



Neville, V., Nakagawa, S., Zidar, J., Paul, E. S., Lagisz, M., Bateson, M., Løvlie, H., & Mendl, M. (2020). Pharmacological manipulations of judgement bias: a systematic review and meta-analysis. *Neuroscience and Biobehavioral Reviews*, 108, 269-286.
<https://doi.org/10.1016/j.neubiorev.2019.11.008>

Peer reviewed version

License (if available):
CC BY

Link to published version (if available):
[10.1016/j.neubiorev.2019.11.008](https://doi.org/10.1016/j.neubiorev.2019.11.008)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Elsevier at <https://www.sciencedirect.com/science/article/pii/S0149763419303252#!>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Pharmacological manipulations of judgement bias: a systematic review and meta-analysis

Vikki Neville^{a,*}, Shinichi Nakagawa^b, Josefina Zidar^c, Elizabeth S. Paul^a,
Malgorzata Lagisz^b, Melissa Bateson^d, Hanne Løvlie^c, Michael Mendl^a

^a*Centre for Behavioural Biology, Bristol Veterinary School, University of Bristol, Langford, BS40 5DU, United Kingdom*

^b*Evolution and Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, New South Wales, Sydney, NSW 2052, Australia*

^c*The Department of Physics, Chemistry and Biology, IFM Biology, Linköping University, SE-581 83 Linköping, Sweden*

^d*Institute of Neuroscience and Centre for Behaviour and Evolution, Newcastle University, Newcastle upon Tyne, NE2 4HH, United Kingdom*

Abstract

Validated measures of animal affect are crucial to research spanning numerous disciplines. Judgement bias, which assesses decision-making under ambiguity, is a promising measure of animal affect. One way of validating this measure is to administer drugs with affect-altering properties in humans to non-human animals and determine whether the predicted judgement biases are observed. We conducted a systematic review and meta-analysis using data from 20 published research articles that use this approach, from which 557 effect sizes were extracted. Pharmacological manipulations overall altered judgement bias at the probe cues as predicted. However, there were several moderating factors including the neurobiological target of the drug, whether the drug induced a relatively positive or negative affective state in humans, dosage, and the presented cue. This may partially reflect interference from adverse effects of the drug which should be considered when interpreting results. Thus, the overall pattern of change in animal judgement bias appears to reflect the affect-altering properties of drugs in humans, and hence may be a valuable measure of animal

*Corresponding author

Email address: vn15961@bristol.ac.uk (Vikki Neville)

affective valence.

Keywords: affective state; animal welfare; judgement bias; meta-analysis; mood disorders; systematic review.

1. Introduction

Measurement of affective state, which is defined as comprising both short-term emotions and longer-term moods [1], and according to dimensional models of affect, valence and arousal components [1], is important to a number of disciplines including psychopharmacology, neuroscience, and animal welfare science, as well as being of societal interest. For example, mood disorders are a significant global concern; it is estimated that 780,000 people died by suicide in 2015, with on average one death every 40 seconds [2]. Major depressive disorder is ranked as the largest single contributor to global disability, and anxiety disorders are ranked sixth [2]. The development of pharmacological treatments for mood disorders has been largely dependent on empirical studies using non-human animals [3, 4]. Reliable and validated measures of affective state in non-human animals are therefore crucial to understanding the neurobiological aetiology of these disorders and to assist in the development of novel treatments. In particular, measures should have both predictive validity (i.e. the extent to which the measure is altered in the predicted direction by drugs which alter human affect) and construct validity (i.e. the extent to which they measure precisely what they claim to measure) [5]. Predictive validity is typically regarded as the 'gold-standard' for validating novel behavioural measures of affective state [6, 7].

Numerous behavioural assays have been developed to assess animal affect. The most common of these include the forced swim test, and its derivative the tail suspension test, which are considered to measure helplessness [8, 9, 10, 11]; the sucrose preference test which is considered to measure hedonic capacity [12, 13]; and the elevated plus maze which is considered to assess the relative value of exploration to safety [14, 15]. Overall, there is good evidence to

suggest that these assays have predictive validity with a broad range of antidepressant or anxiolytic drugs resulting in changes in the predicted direction (i.e. increased latency to immobility; preference for greater sucrose levels; and a greater proportion of time spent in open compared to closed arms of the plus maze). However, when dosed with antidepressant drugs that are used to treat generalised anxiety disorder in humans, rodents do not consistently increase their proportion of time spent in the open compared with closed arms of the open plus maze and hence do not appear to reduce anxiety-like behaviour [16, 17, 18]. Additionally, the construct validity of these assays has been disputed. For example, it has been argued that the forced swim test and tail suspension test reflect a learnt response rather than helplessness [19, 20, 21, 22, 23]. Similarly, research has shown that humans with depression show no reduction in their preference for sucrose over water [24, 25] and that body weight may be a strong confounding factor in the sucrose consumption test [26]. The outlined deficiencies in currently used assays means that there is a clear need for improved methods to measure affective state in non-human animals that have both construct and predictive validity.

The judgement bias task (sometimes referred to as the cognitive bias task or ambiguous cue interpretation task) provides an alternative means to examine affect in non-human animals and has been used widely in the field of animal welfare science since its conception by Harding et al. (2004) [27, 28, 29]. The judgement bias task examines decision-making under ambiguity. Although there is some variation in methodology, the basic principles of the task outlined here are applicable to all judgement bias studies. Individuals are first trained to associate the presentation of one reference cue (e.g. a high frequency tone) with a reward and presentation of another reference cue (e.g. a lower frequency tone) with a lower reward or punisher. Once training is complete, individuals are presented with one or a few untrained probe cues that are intermediate between the reference cues (e.g. medium frequency tones). Their responses to these ambiguous cues are measured to see whether they treat them as signalling the more or less positive outcome. This is measured as latency to

approach the cue or choice to execute or not execute the riskier action which could lead to either the more or less positive outcome (i.e. not the safe action which leads to the null outcome). A decreased latency to approach the cue, or more frequent execution of the riskier action (often deemed 'more optimistic' or 'less pessimistic' in the judgement bias literature), is interpreted to reflect a relatively more positive affective state.

The task is based on the empirical finding that humans experiencing anxiety and depression have a greater expectation of punishing events or reduced expectation of rewarding events than clinically healthy humans [30, 31, 32]. To assess the extent to which judgement bias could measure subjective affective state in humans, the task has been back-translated to human subjects. Studies using the back-translated judgement bias task have demonstrated a correlation between judgement bias and measures of subjectively-experienced affect, such as the State-Trait Anxiety Inventory (STAI), Visual Analogue Scale for Anxiety (VAS-A), and negative affect dimension of the Positive and Negative Affect Schedule (PANAS) [33, 34, 35]. The finding that judgement bias correlates with subjective reports of affective state in humans supports judgement bias as measure of affect, and hence the task appears to have strong construct validity.

A well-established and widely-used approach to validating behavioural measures of affect in non-human animals, which will be used here to assess the validity of judgement bias as a potential measure of affect, is to assess whether drugs with known affect-altering properties in humans produce the predicted shift in the behavioural measure when administered to non-human animals [36, 37, 38, 39]. Specifically, we ask whether pharmacological induction of neurobiological states associated with relatively positive or negative reported affect in humans produces the predicted effect on judgement bias in animals. It is thus important to consider the mechanisms by which such pharmacological manipulations might alter behaviour. There are several cognitive mechanisms that have been proposed to underlie judgement biases and these include changes in attention, perception, reward and punisher sensitivity, prior expectation of rewards and punishers, and action selection (see [40] for

review). Computational analyses of judgement bias data have suggested that both an individual's sensitivity to rewards and punishers and prior expectation of rewards and punishment are key sources of variation in judgement bias [35]. The neurobiological systems underlying these processes and decision-making in general have been subject to much investigation [41, 42, 43, 40]. Briefly, dopaminergic, adrenergic, glutaminergic, and GABAergic activity (particularly in the medial prefrontal cortex, anterior cingulate cortex, thalamus, and locus coeruleus) have been widely implicated in attention and perception [44, 45, 46, 47], serotonergic, dopaminergic, opioidergic, and GABAergic activity (particularly in the orbitofrontal cortex, nucleus accumbens, mesolimbic dopamine projections, and amygdala) have been widely suggested to encode the value and probability of rewards and punishers [48, 49, 40, 43, 50], and adrenergic, serotonergic, and dopaminergic activity (particularly in the basal ganglia and locus coeruleus) have been widely implicated in action selection [51, 52, 53]. Notably, these neurobiological systems are also considered to play a role in human mood disorders (see Box 1). Furthermore, neurotransmitter receptor systems are highly conserved across species [54, 55, 56] and non-pharmacological manipulations designed to induce a positive or negative affective state in non-human animals also result in changes in the activity of these systems [57, 58, 59, 60]. Hence, there are a number of routes by which pharmacologically-induced neurobiological states associated with relatively positive and negative affect (as outlined in Box 1) might alter judgement bias.

Research has been conducted to assess how judgement bias is influenced by affect-altering drugs in non-human animals (See Table 1). Synthesis of these studies would provide an important first step to determine the ability of the judgement bias task to measure pharmacologically-induced neurobiological states associated with positive or negative affect, and hence elucidate the potential validity and reliability of judgement bias as a measure of affect in non-human animals. To this end, we conducted a systematic review and meta-analysis to assess whether pharmacological manipulations alter judge-

ment bias and hence assess the predictive validity of the task. In addition to assessing whether there was an overall effect, we investigated whether the relationship between affect-altering drugs and judgement bias was moderated by factors relating to the drug and administration of the drug, such as the duration and timing of administration, dosage, and neurobiological target of the drug (see Box 1). The potential moderating effects of several task-related factors, such as the presented cue, species used, sex, reinforcement type, response type, and the outcome measure, were also investigated. While we predicted that the effects of judgement bias would be greatest at the ambiguous cues and would depend on dosage, we did not predict that the other moderators would influence the effect of the pharmacological manipulations on judgement bias.

Box 1: Neurobiological targets of affect-altering drugs, i.e. those with antidepressant, anxiolytic, or anxiogenic effects:

Adrenergic system: Epinephrine and norepinephrine are both hormones and neurotransmitters that bind to adrenergic receptors. The adrenergic system is involved in the early stages of a stress response [61]. Brains of depressed patients have reduced levels of norepinephrine and antidepressant drugs such as reboxetine selectively-target the adrenergic system [62, 63].

Dopaminergic system: Dopamine is a neurotransmitter and neuromodulator that can have both inhibitory and excitatory effects on target dopamine neurons. Dopaminergic-system dysregulation is associated with depression [64]. Antidepressant drugs targeting dopamine (although non-specifically) such as monoamine oxidase inhibitors (MAOIs) are available but limited in clinical usage [64].

Gamma-Aminobutyric acid (GABA) system: The neurotransmitter GABA, which binds to GABA receptors, is the major inhibitory neurotransmitter in the brain [65]. Reduced GABA levels are associated with panic disorder [66, 65]. A number of commercially available treatments for anxiety disorders, such as barbiturates, benzodiazepines, and gabapentins, are purported to work by enhancing GABA function.

Glucocorticoid system: Glucocorticoids are a class of steroid hormones that bind to glucocorticoid receptors. The system is involved in the later stages of a stress response, altering cognitive functioning, such as attention and memory, following an acute stressor [61]. Elevated secretion of the glucocorticoid cortisol, specifically upon waking, has been proposed as a biomarker of depression [67, 68].

Glutamatergic system: Glutamate is the brain's major excitatory neurotransmitter and targets glutamatergic receptors that include NMDA, AMPA, and kainite [69]. Several recreational dissociative drugs target the glutamatergic system specifically, such as ketamine and phencyclidine (PCP). NMDA receptor antagonists, such as ketamine, have been found to have antidepressant effects [70, 71].

Opioid system: Opioid receptors are targeted by a number of neuropeptides including endorphins and nociceptin. The opioid system plays a key role in pain modulation, and mediates the euphoric (mood-improving) and analgesic effects of a number of recreational and clinical drugs such as morphine and heroin [72].

Oxytocin system: Oxytocin is a hormone and neuropeptide that targets the oxytocin receptor. The oxytocin system has been implicated in depression; low oxytocin levels have been observed in depressed patients [73].

Serotonergic system: Serotonin is a primarily inhibitory neurotransmitter that binds to serotonergic receptors [48]. There is a wealth of evidence indicating a link between low levels of serotonin and depression [48]. Antidepressant drugs that target the serotonergic system, such as citalopram and fluoxetine, are commonly prescribed [74].

2. Methods

2.1. Ethics statement

Although no animal experiments were conducted directly for the purpose of this meta-analysis, data originating from studies using animals were obtained and analysed. These studies all received ethical approval from the institution at which the research was conducted.

2.2. Literature search

This study followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement (see Fig. 1)[75]. A literature search was first conducted on the 2nd November 2016 to identify all judgement bias studies; the research articles from this literature search were split into groups of those that used pharmacological manipulations (to be analysed here) and those that did not (to be analysed in a separate analysis by Nakagawa et al (*in prep.*)). These meta-analyses were conducted separately as they assessed different research questions; here we specifically want to examine the ability of judgement bias to detect pharmacological manipulations proposed to alter affect and to better understand the factors moderating this, but also due to the complexity the use of different drug doses adds the meta-analysis.

In addition to these articles, a literature search was conducted on the 13th November 2017 using Scopus and Web of Science to identify more recent research papers, and a further literature search was conducted on the 12th July 2019 using Scopus and Web of Science, as well as additional searches in other subject databases (including PsycINFO, PsycARTICLES, PsycBOOKS, PsycEXTRA, PsycTESTS, EMBASE and Medline), grey literature (using ProQuest Dissertation and Thesis Database, Google Data Search and Dimensions platform), and snowballing from reviews on the topic (cited and citing references collected) and from already included papers (citing references collected). Further details on the literature search, including the search-terms used, can be found in the supplementary material.

2.3. Inclusion and exclusion criteria

Following removal of duplicates, the identified articles (Fig. 1) were first screened solely by their abstract. During this abstract-based selection, articles were selected for tentative inclusion in the meta-analysis if they were deemed to be an empirical study which compared judgement bias between at least one control group and at least one treatment group to whom an affect-altering drug had been administered. Additionally, to be included, these studies had to be conducted on vertebrate non-human animals. In this analysis, an affect-altering drug was classified as any substance that was considered to have antidepressant, depressant, anxiolytic, or anxiogenic effects in humans. Where the terms affect-altering, antidepressant, depressant, anxiolytic, or anxiogenic are used throughout this article, they describe the known effect of a drug in human subjects and putative effect of the drug in non-human subjects. Twenty-eight articles met these inclusion criteria.

In the full-text screening, articles were selected on the basis that they had used a variant of Harding *et al*'s (2014) cognitive judgement bias task to compare judgement bias between a group of individuals to whom a vehicle substance had been administered and at least one treatment group who had been given an affect-altering drug [27]. To be included, the outcome measure had to be either latency to approach the cue (e.g. a location in a test arena) on each trial, where approaching the presented cue had been associated with reward and hence shorter latencies would be interpreted as more risk-seeking/less risk-averse behaviour (deemed 'greater optimism' or 'positive judgement bias' within the judgement bias literature), or the proportion of positive responses to each presented cue where a greater proportion would be interpreted as more risk-seeking/less risk-averse behaviour (deemed 'greater optimism' or 'positive judgement bias' within the judgement bias literature), or an outcome measure that could be converted into either form. For example, if the article reported the proportion of negative responses to each presented cue or reported the percentage of positive responses to each presented cue, the extracted data would be subtracted from one or divided by 100 respectively. All included arti-

cles either reported the proportion or latency, but not both, and hence only one of these measures was extracted for each article. Two articles were excluded at the full-text selection stage following retraction by their authors, two were excluded for not meeting the inclusion criteria, and a further two were excluded as they were conference abstracts that duplicated data presented in a journal article which was included in the analysis (Table 1). In addition to these six exclusions, two authors did not provide the requested data and so data from their articles could not be included in the meta-analysis (Table 1). A total of 20 articles were included in the meta-analysis (Table 1).

Fig. 1 here *Table 1 here*

2.4. Data extraction

We extracted the mean and standard deviation of either the latency to approach the presented cue, or proportion of positive responses to the presented cue, as well as the sample size (number of subjects), for every pharmacological treatment and control group for each cue from each article (JZ and VN extracted the data which were checked by VN and SN). Data in a graphical format were extracted using GraphClick 3.0.3 (<http://www.arizona-software.ch/graphclick/>) or WebPlotDigitizer 4.1 (<http://automeris.io/WebPlotDigitizer>). As we extracted mean values, we acknowledge that there may have been variation in how the authors incorporated non-responses into their calculation of latency mean values which we cannot control for. If multiple drug doses had been used, these variables (i.e. mean, standard deviation, and sample size) were extracted for each dosage, and similarly, if there were test sessions that varied the duration between administration and testing or number of days of chronic drug administration, these variables were extracted for each test session. Data collected from both a vehicle and treatment group prior to drug administration were not included as these data did not provide information about the effect of the pharmacological manipulation on judgement bias.

The extracted treatment and control group data were categorised accord-

ing to whether the pharmacological manipulation was expected to induce a neurobiological state associated with either a more or less positively-valenced affective state, which the judgement bias test is predicted to measure. If an anxiogenic or depressant substance had been administered, as determined by the hypotheses stated in the published article alongside the information outlined in Box 1, the treatment group was categorised as the less positive group, and the vehicle group was categorised as the more positive group (i.e. a relatively negative judgement bias was predicted in the treatment group relative to the control group). If an anxiolytic or antidepressant drug had been administered, which was also determined by the hypotheses stated in the article alongside the information outlined in Box 1, the treatment group was categorised as the more positive group and the vehicle group categorised as the less positive group (i.e. a relatively positive judgement bias was predicted in the treatment group relative to the control group). If no hypotheses were stated in the article, this categorisation was based on the description and pharmacodynamics of the substance as outlined on the DrugBank database [76] in addition to the information presented in Box 1. Where multiple doses had been administered, higher doses of anxiolytic or anxiogenic drugs were categorised as more positive whereas higher doses of anxiogenic or depressant drugs were categorised as less positive. This was based on the widespread finding that drugs exert greater effects at higher doses [77, 78].

Information about the article and authors, drug and drug administration, and methodology were also extracted (Table 1 and 2). These included; the article title (Table 1), institute or university at which the research was conducted (extracted but not shown in Table 1 or 2), the name of the drug (Table 2), the dosing duration (chronic - where drugs were administered repeatedly, acute - where the drug was administered immediately before testing, or chronic wash-out - the period after drug administration had stopped following chronic administration), the time between administration and testing (acute studies only), the number of days since the first dose (chronic studies only), the dosage (in mg/kg), the neurobiological target of the drug, the pharmacological manipu-

lation type (antidepressant/anxiolytic or depressant/anxiogenic), the species tested, and the outcome variable used (latency or proportion), cue (positive reference cue, midpoint probe cue, negative reference cue, and where included the near negative probe cue and near positive probe cue) (not shown in Table 1 or 2, although number of probe cues given instead), sex of the experimental subjects (all male, all female, or both male and female), reinforcement type used for the reference cue (reward-punishment - where the positive reference cue was rewarded and negative reference cue punished; reward-null - where the positive reference cue was rewarded and negative reference cue was not rewarded; or reward-reward - where the positive reference cue was rewarded with a high reward and negative reference cue was reward with a low reward), response type which reflected whether both or only one of the reference cues required an approach response (go/no-go - where the positive reference cue required an active response and the negative reference cue required no response, or go/go - where both reference cues required an active response), the proportion of probe trials in relation to the total number of trials, and cue type (reference or probe) (Table 2). To ensure that dosage was comparable between substances and species, each drug dose within a species was standardized by dividing the dosage (in mg/kg) by the standard deviation of all doses administered within each drug for each species.

Table 2 here

2.5. *Effect size and sampling variance calculation*

As latency data are bounded at zero and proportion data are bounded between zero and one data obtained from the judgement bias task do not follow a Gaussian or normal distribution. The delta method (Taylor approximation) was used to adjust the extracted mean (\bar{x}) and (sampling) variance (sd^2) prior to calculating the effect size to account for the non-normality of the raw data [79]. For extracted latency data, which were assumed to follow a log-normal distribution, this adjustment was calculated via the log transformation as:

$$\overline{\ln(x)} = \ln(\bar{x}) - \ln\left(\sqrt{1 + \frac{sd^2}{\bar{x}^2}}\right) \quad (1)$$

$$sd_{\ln}^2 = \ln\left(1 + \frac{sd^2}{\bar{x}^2}\right) \quad (2)$$

In this case, the transformed sampling variance is exact and not an approximation.

For extracted proportion data, which were assumed to follow a binomial distribution, this adjustment was calculated via the logit transformation as [80]:

$$\overline{\text{logit}(x)} = \text{logit}(\bar{x}) + \frac{sd^2}{2} \left(\frac{1}{(1 - \bar{x})^2} - \frac{1}{\bar{x}^2} \right) \quad (3)$$

$$sd_{\text{logit}}^2 = sd^2 \left(\frac{1}{\bar{x}} + \frac{1}{1 - \bar{x}} \right)^2 \quad (4)$$

Hedge's g [81], a measure of effect size based standardized differences in means, was then calculated as the difference between the means of the relatively positive treatment (in which a relatively more positive affective state was expected, as outlined above) \bar{x}_{+ve} and means of the relatively negative treatment (in which a relatively less positive affective state was expected, as outlined above) \bar{x}_{-ve} , divided by the pooled standard deviation, sd_{pool} , and then adjusted for biases arising from small sample sizes by factor J which depended on the sample size of the relatively positive n_{+ve} and relatively negative n_{-ve} groups:

$$\text{SMD} = \frac{\bar{x}_{+ve} - \bar{x}_{-ve} \cdot J}{sd_{pool}} \quad (5)$$

$$(6)$$

$$sd_{pool} = \sqrt{\frac{(n_{+ve} - 1)sd_{+ve}^2 + (n_{-ve} - 1)sd_{-ve}^2}{n_{+ve} + n_{-ve} - 2}} \quad (7)$$

$$(8)$$

$$J = \left(1 - \frac{3}{4(n_{+ve} + n_{-ve}) - 9} \right) \quad (9)$$

For the latency data, Hedge's g was multiplied by minus one to account for a higher proportion being equivalent to a lower latency, in terms of judgement bias.

The sampling variance was calculated as follows:

$$se_{\text{SMD}}^2 = \frac{n_{+ve} + n_{-ve}}{n_{+ve}n_{-ve}} + \frac{\text{SMD}^2}{2(n_{+ve} + n_{-ve})} \quad (10)$$

To account for shared controls, if one vehicle treatment group was compared to multiple drug treatment groups, an additional effect size and sampling variance was calculated based on a sample size for the vehicle group that had been divided by the number of treatment groups [82].

2.6. Meta-analysis and meta-regression models

The meta-analysis and meta-regression were conducted using the function, `rma.mv` from the R [83] package `metafor` [84]; this function allowed us to fit multilevel meta-analytic and meta-regression models [85]. All models included drug, institution at which the research was conducted, and effect ID (a unique ID given to each effect size) as random effects to account for the non-independence of effect sizes from studies conducted at the same institute or using the same drug [86], and were fit using restricted maximum likelihood. The Knapp and Hartung adjustment was applied to all analyses [87]. Initially, an intercept only model was fit to the effect sizes. A p -value for this model was obtained using a Wald-type test based on a t -distribution. Heterogeneity was assessed by calculating the I^2 values for each random effect in the model and an overall I^2 value for the model, following [88], which is an extension of the original I^2 [89].

Meta-regression was used to examine whether the following categorical and continuous moderators significantly contributed to variation between effect sizes: the dosing duration (chronic, acute, or chronic wash-out), the time between administration and testing (acute studies only), the number of days since the first dose (chronic studies only), the dosage differences between treatments from which the effect size was calculated, the neurobiological target of

the drug, the manipulation type (positive or negative affect induction), the species tested, and the outcome variable used (latency or proportion), presented cue (positive reference cue, near-positive probe cue, midpoint probe cue, near-negative probe cue, negative reference cue), sex of the experimental subjects (all male, all female, or both male and female), reinforcement type (reward-punishment, reward-null, or reward-reward), response type (go/no-go, or go/go), cue type (reference or probe), and proportion of probe cues to reference cues in the test session. An omnibus test based on an F distribution, which examines the degree of variance explained by a moderator, was used to assess the significance of each moderator [90]. To further investigate significant moderators, pairwise comparisons were made between the mean effect size for each level of the moderator. A Wald-type test was used to assess the significance of these pairwise comparisons. Moderators which were significant in the meta-regression were subsequently included together in a full model and their influence on the effect sizes was re-assessed. To verify that the model of best fit included all moderators, Akaike's information criterion (AIC) was calculated for the full model and was compared to models where a moderator had been removed.

2.7. *Subset analyses*

As affect is hypothesised to exert a greater influence on decision-making under ambiguity than under certainty, any treatment designed to pharmacologically induce a neurobiological state associated with a relatively more positive or negative affective state is expected to have the greatest influence on judgement bias at the ambiguous probe cues (see Fig. 2 for example of hypothesised data) [40, 1]. There are also methodological and theoretical reasons as to why an effect may be observed at one cue and not others. For example, a cue may be too perceptually similar to either of the reference cues for there to be ambiguity about the outcome, or a potential punisher may be much more aversive than the reward is rewarding, to the extent that all animals will avoid probe cues that are similar to the negative reference cue. By considering all

cues equally (including reference cues), the effect of an affective manipulation might be obscured, potentially leading to the false inference of no significant effect. To this end, we conducted an additional analysis on a subset of data that included only the effect sizes from the probe cue with the largest absolute effect size for each drug within an article. Additionally, we analysed a second subset of data that included only the effect sizes for the cue with the absolute largest effect size in the direction of the mean effect size for each drug within an article to avoid including outlying effects that might not necessarily reflect the influence of the manipulation. If only one probe cue was presented in a study, data from this probe cue were included in the subset data.

Fig. 2 here

2.8. Publication bias and sensitivity analysis

To assess the reliability of results across different analytical approaches and to check for a publication bias, the intercept-only and full meta-regression model were re-fit to the data under a Bayesian statistical framework using the R package MCMCglmm [91]. The non-independence of effect sizes can also be accounted for using Bayesian methods. A parameter-expanded prior, allowing variance components to have different prior distributions, was used for both the random effect of drug and institution ID, while the prior variance for random effect of effect ID was fixed at one. Model fitting had 110,000 iterations, 10,000 burn-in periods, and thinning by every 100, resulting in an effective sample size of 1000. The result of this intercept-only model was compared to our initial intercept-only model. The 'meta-analytic residuals' (*sensu* [88]) from full meta-regression model conducted in MCMCglmm were used to produce a funnel plot and run Egger's regression, which here regresses the meta-analytic residuals against precision [92, 88], and hence checks for a publication bias. Additionally, the intercept-only meta-analysis was repeated but with the effect size and sampling variance that had been adjusted (via the sample size) for shared controls, to assess whether this altered the results.

3. Results

3.1. Data review

We extracted 557 effect sizes from 20 articles that had been published by authors based at 10 different institutions (see Table 1 and 2). Twenty-seven different drugs were used across these studies. The majority (328) of the effect sizes came from studies that had used drugs expected to induce a relatively positive affective state (anxiolytics or antidepressants, 12 articles), while the remainder used anxiogenic or depressant drugs (112 effect sizes, 9 articles). There were 408 effect sizes (14 articles) that came from studies using acute pharmacological manipulations, 97 effect sizes (6 articles) from studies using chronic pharmacological manipulations, and 52 effect sizes (5 articles) that came from the wash-out period of a chronic pharmacological manipulation. Most effect sizes came from studies using drugs that targeted the serotonergic system (198 effect sizes, 7 articles) while a high proportion of studies also used drugs that targeted a range of neurobiological systems (190 effect sizes, 9 articles) which included drugs such as cocaine and d-amphetamine which target the dopaminergic, serotonergic, and adrenal systems. The remaining effect sizes were from experiments using drugs that specifically targeted GABAergic system (46 effect sizes, 4 articles), adrenergic system (43 effect sizes, 3 articles), dopaminergic system (36 effect sizes, 1 article), opioid system (20 effect sizes, 1 article), glucocorticoid system (15 effect sizes, 1 article), oxytocin system (9 effect sizes, 2 articles). Five different species were used across the studies; the most frequently used species according to the number of effect sizes was rat (418 effect sizes, 11 articles), followed by pig (60 effect sizes, 2 articles), sheep (50 effect sizes, 4 articles), chicken (23 effect sizes, 2 articles), and dog (6 effect sizes, 1 article). Proportion was more commonly used as the outcome measure (435 effect sizes, 12 articles) compared with latency (122 effect sizes, 8 articles). The majority of effect sizes came from studies using only male subjects (421 effect sizes, 11 articles), followed by only female subjects (130 effect sizes, 8 articles), and six effect sizes (1 articles) came from studies that used both male and fe-

male subjects. The most common reinforcement type was reward-punisher (420 effect sizes, 16 articles), followed by reward-reward reinforcement (131 effect sizes, 3 articles), and reward-null (6 effect sizes, 1 article). There were more effect sizes from studies using a 'go/go' design (415 effect sizes, 10 articles) compared with a 'go/no-go' design (142 effect sizes, 10 articles).

Across the articles from the acute studies, the average time between the administration of the drug and testing was 32.903 ± 5.530 (mean \pm SE) minutes. The average number of days between the start of the chronic drug treatment and testing was 9.000 ± 1.074 , and the average days the animal had been withdrawn from a drug when tested in the wash-out period was 6.938 ± 0.824 . The mean proportion of probe cues to reference cues used during a test session was 0.341 ± 0.037 . There were 11 articles that used more than one probe cue and three of these articles examined the effect of more than one drug. In total, there were 14 sets of effect sizes obtained from different articles using different drugs which used more than one probe cue. The probe cue with the greatest absolute effect size was the near-positive probe cue on nine occasions, the near-negative probe cue on four occasions, and the midpoint probe cue on one occasion. The probe cue with the greatest absolute effect size was also the presented cue with the greatest absolute effect size in the direction of the mean effect for all but one of the sets of effect sizes, where the near-positive probe cue had the greatest absolute effect sizes and the near-negative probe cue had the greatest absolute effect size in the direction of the mean effect.

3.2. Meta-analysis

Overall, considering all effect sizes equally, affect-altering drugs did not significantly induce a judgement bias in non-human animals, although a small effect size (*sensu* [93]: small=0.20, moderate=0.5, large=0.8) was observed (mean=0.239, 95% confidence interval or CI=-0.047-0.525, $t_{556}=1.639$, $p=0.102$). However, this needs to be interpreted in the context of the observed high total heterogeneity in the model, with an I^2 value of 89.535 (>75%=high, [89]), indicative of wide variation in the extent to which pharmacological manipula-

tions alter judgement bias that warrants further examination. The between-effect-size effect (i.e., residuals; 42.378%) and the between-drug effect (i.e., which drug were used; 35.112%) explained a large percentage of this heterogeneity, while a smaller percentage of variability was due to institutional variation (12.044%). Heterogeneity between effect sizes was further explored through the meta-regression.

3.3. *Subset analyses*

However, as aforementioned, given the theoretical framework for judgement bias, we did not anticipate that effect sizes would be equal across all cues. Instead, we considered it likely that pharmacological manipulations would exert the greatest influence at only one of the probe cues, with proximity of this cue to the reference cues differing between studies as a result of different methodologies. Hence, subset analyses were conducted to assess the extent to which the pharmacological manipulations of affect altered judgement bias at, at least, one of the probe cues. Pharmacological manipulations using drugs which have affect-altering properties were found to have a significant small to moderate effect on judgement bias when the analysis was repeated on the subset data comprising only data from the ambiguous cue with the largest absolute effect size (mean=0.394, CI=-0.046-0.727, $t_{154}=2.067$, $p=0.040$), and a significant small to moderate effect on the subset data comprising only data from the ambiguous cue with the largest absolute effect size in the direction of the mean effect (mean=0.455, CI=0.061-0.849, $t_{154}=2.279$, $p=0.024$).

3.4. *Meta-regression*

The meta-regression revealed that several moderators significantly explained the observed heterogeneity among the extracted effect sizes, these moderators were: pharmacological manipulation type (Fig. 3: $F_{(1,555)}=16.056$, $p<0.001$), dosage ($F_{(1,519)}=6.614$, $p=0.010$), the reinforcement type ($F_{(2,554)}=3.653$, $p=0.027$) and the cue type ($F_{(1,555)}=4.725$, $p=0.030$). The presented cue ($F_{(4,552)}=2.002$, $p=0.093$), and the neurobiological target of the

drug were marginally non-significant moderators ($F_{(8,548)}=1.835$, $p=0.079$). More specifically, pharmacological manipulations expected to induce a relatively negative affective state (either depressant or anxiogenic) had a greater effect on judgement bias than those expected to induce a relatively positive affective state (Table 3). Greater differences in dosage between the relatively positive and negative treatments were associated with smaller effect sizes. The greatest effect size was found when the reinforcement used for the reference cues was a high reward and low reward, compared with a reward and punisher. The effect of the pharmacological manipulation was greater at the probe cues compared with the reference cues (Table 3). The effect of the pharmacological manipulation was weaker at the positive reference cue compared to the midpoint probe cue, and near-positive probe cue, and tended to be weaker than the negative reference cue (Table 3). There was no difference in effect size at the positive reference cue compared with the near-negative reference cue (Table 3). The remaining moderators tested were not found to significantly explain variation in effect size. The effect of drugs targeting the adrenergic system differed significantly from all other drugs used apart from drugs targeting the opioid and oxytocin system (Table 3). Drugs targeting the adrenergic system had the opposite effect than expected; a negative judgement bias was induced when a positive judgement bias was hypothesised. Other moderators with non-significant effects included: species ($F_{(4,552)}=0.835$, $p=0.503$), dosing frequency ($F_{(2,554)}=0.108$, $p=0.898$), time since last dose ($F_{(1,406)}=0.467$, $p=0.495$), number of days since first treatment ($F_{(1,95)}=1.169$, $p=0.282$), sex ($F_{(2,554)}=0.328$, $p=0.720$), response type ($F_{(1,555)}=0.040$, $p=0.842$), proportion of ambiguous cues to reference cues ($F_{(1,555)}=1.531$, $p=0.217$), and outcome measure ($F_{(1,555)}=0.139$, $p=0.709$).

Figure 3 here *Table 3 here*

The best fitting model included cue (i.e. positive reference cue, midpoint probe cue, negative reference cue, near negative probe cue and near positive probe cue) instead of cue type (i.e. reference or probe) (ΔAIC (i.e. difference in AIC values between models)=0.437), and all significant moderators

identified in the univariate meta-regression. Removal of the neurobiological target of the drug ($\Delta\text{AIC}=13.511$), dosage ($\Delta\text{AIC}=6.705$), cue ($\Delta\text{AIC}=3.583$), reinforcement type ($\Delta\text{AIC}=7.573$), and manipulation type ($\Delta\text{AIC}=6.465$) resulted in a poorer fit according to the AIC values. The best fitting model had a marginal R^2 value (*sensu* [94]) of 72.844%. In this model, the difference between effect sizes where a relatively positive compared with relatively negative affective state had been induced was significant, with a moderate effect size ($\Delta\text{mean}=0.582$, $\text{CI}=0.054\text{-}1.110$, $t_{506}=2.164$, $p=0.031$). Effect sizes from drugs targeting the adrenergic system were overall in the opposite direction to expected and there was a large and significant difference in effect sizes between adrenergic system targeting drugs and multiple system targeting drugs ($\Delta\text{mean}=0.852$, $\text{CI}=0.043\text{-}1.661$, $t_{507}=2.069$, $p=0.039$) and GABAergic system targeting drugs ($\Delta\text{mean}=1.299$, $\text{CI}=0.369\text{-}2.228$, $t_{507}=2.746$, $p=0.006$), and a large but marginally non-significant difference in effect sizes between adrenergic system targeting drugs and serotonergic system targeting drugs ($\Delta\text{mean}=0.817$, $\text{CI}=-0.073\text{-}1.707$, $t_{507}=1.803$, $p=0.072$), dopaminergic system targeting drugs ($\Delta\text{mean}=0.936$, $\text{CI}=-0.083\text{-}1.956$, $t_{507}=1.804$, $p=0.072$), glucocorticoid system targeting drugs ($\Delta\text{mean}=1.451$, $\text{CI}=-0.250\text{-}3.151$, $t_{507}=1.676$, $p=0.094$). There was a large but non-significant difference between the effect sizes of drugs targeting the adrenergic and oxytocin system ($\Delta\text{mean}=0.961$, $\text{CI}=-0.894\text{-}2.815$, $t_{507}=1.018$, $p=0.309$) and moderate but non-significant difference between the effect sizes of drugs targeting the adrenergic compared with opioid targeting drugs ($\Delta\text{mean}=0.555$, $\text{CI}=-0.621\text{-}1.732$, $t_{507}=0.927$, $p=0.354$). Effect sizes were significantly weaker at the positive reference cue compared with the midpoint probe cue ($\Delta\text{mean}=0.163$, $\text{CI}=0.019\text{-}0.308$, $t_{507}=2.218$, $p=0.027$), and near-positive probe cue ($\Delta\text{mean}=0.289$, $\text{CI}=0.025\text{-}0.553$, $t_{507}=2.148$, $p=0.032$). Effect sizes at the positive reference cue were not significantly different from effect sizes at the negative reference cue ($\Delta\text{mean}=0.118$, $\text{CI}=-0.028\text{-}0.264$, $t_{507}=1.592$, $p=0.112$) or at the near-negative probe cue ($\Delta\text{mean}=0.206$, $\text{CI}=-0.060\text{-}0.472$, $t_{507}=1.519$, $p=0.129$). There was a small and significant difference between effect sizes from studies using

high and low rewards as the more and less favourable outcome, respectively, compared with studies which used rewards and punishers ($\Delta\text{mean}=-0.366$, $\text{CI}=-0.626-0.106$, $t_{543}=-2.764$ $p=0.006$), while there was a moderate but non-significant difference between studies that used high and low rewards and those that used a reward and null outcome ($\Delta\text{mean}=-0.546$, $\text{CI}=-1.548-0.457$, $t_{543}=-1.070$, $p=0.285$).

3.5. *Study exclusion*

As the initial analysis revealed that drugs targeting the adrenergic system had the opposite effect on judgement bias than hypothesised, which differed significantly from the majority of the other drugs not specifically targeting the adrenergic system demonstrating that these drugs produced an anomalous effect in comparison to all other drugs examined, we re-assessed the rationale for classification of adrenergic-targeting drugs as either anxiolytic/antidepressant or anxiogenic/depressant. There is conflicting evidence about the affect altering properties of adrenergic system targeting drugs in both non-human [95, 96] and human [97] animals. Moreover, studies have demonstrated that acute negative affect is associated with increased levels of norepinephrine, while chronic negative affect is associated with decreased levels of norepinephrine [98, 99], which further complicates assessment of whether the neurobiological state induced by the pharmacological manipulations most resembled that of a relatively negative or positive affective state. Consequently, we made the post-hoc decision to re-analyse the data excluding effect sizes from studies using adrenergic-system targeting drugs. Three studies had used adrenergic-system targeting drugs; one study had used clonidine and the other two studies had used reboxetine. Both clonidine and reboxetine are considered to induce a relatively positive affective state. These studies accounted for 7.719% (43) of the effect sizes analysed.

3.6. *Post-exclusion meta-analysis*

Following the exclusion of these effect sizes, a moderate overall effect was observed; pharmacological manipulations which induced a neurobiological

state associated with relatively more positive or negative affect were found to significantly influence judgement bias in the predicted direction (mean=0.400, CI=0.056-0.744, $t_{513}=2.287$, $p=0.023$). However, there again existed high heterogeneity ($I^2=89.746\%$); with 38.362% attributable to between-effect-size effects, 20.732% to between-drug effects, and 30.653% to institutional variation. The meta-analysis using both data subsets (using only one probe cue) revealed a significant and moderate overall effect of pharmacological manipulations on judgement bias (absolute greatest probe cue effect sizes: mean=0.520, CI=0.116-0.924, $t_{144}=2.543$, $p=0.012$; and absolute greatest probe cue effect in direction of mean: mean=0.579, CI=0.157-1.001, $t_{144}=2.711$, $p=0.008$).

3.7. Post-exclusion meta-regression

While manipulation type (Fig. 4: $F_{(1,512)}=15.700$, $p<0.001$) and dose (Fig. 4: $F_{(1,476)}=5.169$, $p=0.023$), remained significant as moderators when studies using adrenergic system targeting drugs were excluded, the presented cue (Fig. 4: $F_{(4,509)}=2.396$, $p=0.049$) was now significant as opposed to marginally non-significant and drug target ($F_{(6,506)}=0.578$, $p=0.748$) cue type ($F_{(1,512)}=2.594$, $p=0.108$), and reinforcement type ($F_{(2,511)}=0.144$, $p=0.866$) were no longer significant. The model which included all three significant moderators provided a better fit than the models which excluded manipulation type ($\Delta AIC=13.465$), cue ($\Delta AIC=2.299$), and dose ($\Delta AIC=5.430$). This full model had a marginal R^2 value of 61.008%.

The difference between effect sizes at the midpoint probe cue and positive reference cue was very small but significant ($\Delta mean=0.154$, CI=0.011-0.297, $t_{471}=2.111$, $p=0.035$, and the difference between effect sizes at the positive reference cue and near-positive probe cue was small but marginally non-significant ($\Delta mean=0.245$, CI=-0.035-0.525, $t_{471}=1.716$, $p=0.087$). Contrary to the previous analysis including adrenergic-targeting drugs, a very small but significant difference was found between effect sizes at the negative and positive reference cues, with greater effect sizes at the negative reference cue ($\Delta mean=0.183$, CI=0.039-0.327, $t_{471}=2.491$, $p=0.013$). The difference in effect sizes between

the positive reference and near-negative probe cue remained non-significant ($\Delta\text{mean}=0.192$, $\text{CI}=-0.089-0.474$, $t_{471}=1.341$, $p=0.181$). Effect sizes were still observed to be greater when the anxiogenic or depressant drugs were used compared to the antidepressant or anxiolytic drugs with a moderate difference in effect sizes ($\Delta\text{mean}=0.701$, $\text{CI}=0.348-1.055$, $t_{471}=3.897$, $p<0.001$), and effect sizes remained significantly greater when there were smaller differences in dosage between the relatively positive and relatively negative treatment, although the effect was very small ($\text{mean}=-0.0379$, $\text{CI}=-0.071-0.004$, $t_{471}=-2.223$, $p=0.027$).

Figure 4 here

3.8. Publication bias and sensitivity analysis

The results of the Bayesian meta-analysis were consistent with the results of our likelihood-based meta-analyses both prior to and following the removal of effect sizes from studies using drugs targeting the adrenergic system. The effect of the pharmacological manipulations on judgement bias was not significant prior to data exclusion ($\text{mean}=0.242$, 95% credible interval= $-0.097-0.666$, $p=0.194$), but a marginally non-significant overall effect emerged following data exclusion from studies using adrenergic-system targeting drugs ($\text{mean}=0.387$, credible interval= $0.020-0.864$, $p=0.056$).

Visual inspection of the funnel plots produced from the meta-analytic residuals and raw effect sizes (Fig. 5) did not indicate that a publication bias was present, nor did the results of Egger's test on either the analysis prior to ($t_{519}=-0.419$, $p=0.675$) or following ($t_{476}=0.568$, $p=0.570$) the exclusion of data.

Figure 5 here

Re-analysis of the intercept-only model using the effect sizes and variances that had been adjusted for shared controls did not alter the results qualitatively. The result prior to data exclusion was statistically non-significant ($\text{mean}=0.240$, $\text{CI}=-0.047-0.527$, $t_{522}=1.641$, $p=0.101$) while following data exclusion was significant ($\text{mean}=0.401$, $\text{CI}=0.057-0.745$, $t_{476}=2.291$, $p=0.022$).

4. Discussion

Judgement bias is a relatively new and promising measure of animal affect that may provide a useful alternative to more common behavioural assays used to assess the efficacy of potential pharmacological treatments of mood disorders, such as the forced swim test. Empirical studies with human subjects have supported its construct validity [30, 31, 32, 33, 34]. To examine its predictive validity, we conducted a systematic review and meta-analysis of studies investigating the effect of affect-altering drugs on judgement bias in non-human animals. We analysed data from 20 published research articles which yielded 557 effect sizes.

There was high heterogeneity ($>75\%$) between the effect sizes observed [100] indicating strong variability in the extent to which pharmacological manipulations of affective state alter judgement bias. The drug used accounted for some of this variability, as did the institution at which the research was conducted, yet a high proportion of heterogeneity was also attributed to variation within drug and institution. Our meta-regression further highlighted a number of factors which explained variation in effect sizes including the neurobiological drug target, manipulation type (whether the drug was hypothesised to induce a negative or positive affective state), dosage, cue, and cue type (reference or probe).

Initially, considering all effect sizes across all cues equally (including reference cues), we found no significant overall effect of affect-altering drugs on judgement bias in non-human animals. However, because there are theoretical and empirical reasons for an effect being more likely at ambiguous cues as opposed to reference cues, and/or to occur at one ambiguous cue but not others (e.g. because the others may happen to be too perceptually similar to the reference cues, see Methods), considering all cues equally may obscure an effect of a treatment manipulation. Indeed, judgement bias studies often report effects that are observed at only a subset of (ambiguous) cues (e.g. [101]). Consequently, we also carried out analyses using subsets of data that included

only (i) effect sizes for the probe cue with the largest absolute effect size; (ii) effect sizes for the cue with the absolute largest effect size in the direction of the mean effect size. These analyses revealed that the pharmacological manipulations altered judgement bias in the predicted direction.

The results of the meta-regression showed a clear moderating effect of the neurobiological drug target, particularly of drugs targeting the adrenergic system, whose effect differed significantly from the majority of other drugs used. A small to medium effect using data from all cues was found following the removal of data from studies targeting the adrenergic system, and a moderate effect was found when considering data from the subset analyses described above. Thus, this meta-analysis provides support for the validity of judgement bias as measure of affect in non-human animals, having demonstrated that pharmacological manipulations using drugs known to influence affect in humans overall alter judgement bias in non-human animals at the probe cues in the predicted direction. However, there exist a few caveats; it is important to state that this result should not be interpreted as evidence that the pharmacological manipulations did alter affect and that this shift in affect directly influenced judgement bias. Instead, this result demonstrates that pharmacologically-induced neurobiological states associated with relatively positive or negative affect alter judgement bias in the predicted direction, given hypotheses about how affect should alter judgement bias - judgement bias reliably predicted the affect-altering properties of the pharmacological manipulation. However, the pharmacological manipulations may have altered judgement bias through a variety of mechanisms and we cannot preclude the possibility that manipulations did not alter affect in non-human animals as they would in human subjects, or that the observed shift in judgement bias was directly attributable to a shift in affect.

The three excluded studies used either reboxetine, an antidepressant, or clonidine, which is used off-licence to treat anxiety disorders. Jointly, these drugs were found to exert an opposite effect on judgement bias; inducing a negative judgement bias when a positive judgement bias was predicted. Para-

doxically, depression and anxiety are known side effects of these drugs [102]. Moreover, these studies both used an acute dose which may explain why their effects were not in the predicted direction. Both norepinephrine and cortisol increase in response to stress and acute dosing of drugs which simultaneously elevate levels of norepinephrine and cortisol have been shown to result in stress-like changes in the neural response to negative stimuli in humans [103]. It is therefore feasible that the acute delivery of adrenergic-system targeting antidepressant drugs induced a neurobiological state associated with relatively negative rather than a positive affective state which resulted in the relatively negative judgement bias observed. This potential explanation is further supported by studies that have observed anxiety-like states in rodents following the administration of similar adrenergic-system targeting drugs [95, 96].

An alternative explanation could be related to another side effect of adrenergic agonists that has been documented in human and non-human animal subjects; sedation [104, 102]. A sedated animal may not have been able to fully partake in the experiment or have been considerably slower to respond, leading to seemingly risk-averse (deemed 'pessimistic' in the judgement bias literature) responses. This is perhaps further supported by the finding that clonidine and reboxetine led to an increased latency to respond to the positive reference cue, as described by the authors of the studies included in this meta-analysis. However, further studies would be required to reveal the extent to which the results from these two studies can be generalised to all adrenergic system targeting drugs.

Depressant and anxiogenic drugs, had a greater effect on measured judgement bias than antidepressant and anxiolytic drugs. This result may reflect an interaction between the drugs administered and affective states arising from the process of being tested, which may sometimes be negative in their own right (e.g. invasive administration of drugs, social isolation during testing, and potential delivery of an aversive decision outcome). These factors may have enhanced the neurobiological effect of the depressant and anxiogenic drugs, while dampening the neurobiological effect of the antidepressant and anxi-

olytic drugs. Notably, the potential negative affective state induced by testing may also explain the finding, prior to exclusion of adrenergic targeting drugs, that effect sizes were greater when only rewards were used as reinforcement, as opposed to both reward and punishers. Indeed, in humans there is evidence to suggest that some affect-altering recreational drugs intensify the affective state of an individual prior to consumption, or result in the exaggerated interpretation of emotional stimuli [105, 106]. With regards to the development of novel treatments for mood disorders such as depression and anxiety, this perhaps suggests that attention should be given to the potential effects of the testing procedure on affect, and that greater sample sizes may be required to provide sufficient power for an effect of the potential pharmacological treatment to be detected.

Another possible explanation for the moderating effect of manipulation type is that there are floor effects which limit the impact antidepressant and anxiolytic drugs may have on judgement bias. There will be a physical limit to how quickly an animal can approach a cue, and the control animals may already be performing at or close to this limit, meaning that the animals that had been given a antidepressant and anxiolytic drug could not respond any quicker. However, this explanation will only be relevant to studies measuring approach latency. Similarly, the smoke-detector principle states that individuals should be overly responsive to potential threats [107]; just as the cost of a smoke detector not detecting a fire is far greater than the cost of the smoke detector sounding an alarm when there is no fire, false positives are also optimal in the detection of predators. An individual may continue to appear relatively cautious even when in a more positive affective state because the cost of not avoiding punishers (i.e. potential death) is so high that it would be suboptimal for an individual to behave in a more risky manner (i.e. making the 'optimistic' response which could lead to a punisher, as opposed to the 'pessimistic' response which is the safe option) [107, 108].

Although this meta-analysis did not identify a difference in effect sizes between studies which analysed proportion or latency data, it is important to

discuss these two outcome measures. Typically, both latency and proportion data can be collected from judgement bias studies and it is unclear what drives a researcher to select either measure. There are merits and disadvantages of using either measure; while latency contains more information than proportion data, in the sense that as a continuous variable it may identify variation that proportion data cannot, it may also be more subject to influences from other factors such as any effect of the drug on motor responses as outlined above or other cognitive biases such as attention biases [40].

Greater effects were observed when there were relatively smaller differences in dosage between treatments. This is consistent with the inverted U-shaped dose-response function that is sometimes observed in drug studies [109, 110, 111]. This result may reflect that higher doses increase the probability of side effects which may interfere with task performance [112, 113]. Adverse effects that alter the motivation of the animal (e.g. reduced appetite), their consummatory behaviour (e.g. nausea), or psychomotor abilities (e.g. sedation) are likely to affect judgement bias. Such side effects are common to several affect-altering drugs [102]. It may therefore be sensible to take measures of activity or food consumption concurrent to the judgement bias task to assess the potential impact of side effects of drug manipulations.

The meta-analysis also found that the effects of the pharmacological manipulations on judgement bias were weakest at the positive reference cue, and that the effect of pharmacological manipulations was greater when only the probe cue with the greatest effect size within each drug and article were analysed. This reflects that pharmacological manipulations using affect-altering drugs exert a stronger influence on trials where there is ambiguity about the outcome of the trial compared with trials where the reward is certain. On presentation of the positive reference cue, there should be little ambiguity about the outcome, and it would be expected that the animal should make the response that allows them to obtain the reward on a high proportion of trials. The influence of any manipulation which putatively alters affect should be greatest when there is uncertainty about the outcome as subjective probabil-

ities of uncertain outcomes are thought to be more strongly informed by an individual's affective state [1, 114, 115]. Thus, this finding is consistent with the theoretical framework underlying judgement bias. However, given that a number of cognitive processes could lead to a shift in judgement bias at the probe but not the positive reference cue, this finding does not negate the possibility that cognitive processes other than probability estimation underlie the relationship between the pharmacological manipulations and decision-making on the judgement bias task.

It is unclear why the extracted effect sizes were not smaller at the negative reference cue following exclusion of effect sizes from studies that had used adrenergic-targeting drugs. The pharmacological manipulations were not expected to exert a similar effect at the negative reference cue compared with the probe cues, as there should be little uncertainty about the outcome when the reference cues are presented. Moreover, in studies in which multiple probe cues were presented, the pharmacological manipulations rarely exerted the greatest influence at midpoint cue, where there should be the greatest uncertainty about the outcome. This further suggests that pharmacologically-induced neurobiological states associated with relatively more positive or negative affect do not necessarily induce a greater judgement bias as uncertainty about the decision outcome increases. Both valuation and probability of decision outcomes have been identified as key components of decision-making that underlie variation in the judgement bias task [35]; an individual might be more likely to make a risky or more 'optimistic' response if they considered the reward to more probable or punisher to be less probable or if they considered the reward to be more valuable or punisher to be less aversive [40, 43]. Speculatively, it is possible that the pharmacological manipulations altered the valuation of the punisher, hence altering responses to its presentation. A better understanding of the cognitive processes underlying judgement bias and how they relate to affect, which could be achieved through a battery of tests or computational modelling of judgement bias data, would be highly valuable. Similarly, it is possible that some findings reflect differences between the re-

ward and punisher systems. Indeed, prior to the removal of adrenergic drugs from the dataset, we found that the effect of the pharmacological manipulation was greater when no reward was used as the negative outcome compared to when a punisher was used as the negative outcome.

Our meta-analysis found no evidence to indicate that the species used, the dosing frequency, the time since last dose in acute studies, the number of days since first treatment in chronic studies, the outcome variable used, the biological sex of the individuals studied, the reinforcement type, or response type had moderating effects on the influence of pharmacological manipulations of affect on judgement bias. While this might reflect that there is insufficient power to detect an effect, it might indicate that judgement bias is robust to variation in methodology and across species. Interestingly, despite being one of the most commonly used non-human animal species in research, none of the studies included in this meta-analysis used mice [116, 117]. As judgement bias tasks have been successfully conducted in mice [118, 119], we consider that it would be highly worthwhile to examine the extent to which pharmacological manipulations alter judgement bias in mice. We found no evidence to suggest a publication bias.

Future studies should attempt to account for the potential side effects of pharmacological manipulations. Observing behaviour following drug administration, for example activity levels and food and water consumption, may help to highlight potential adverse effects. The majority of effect sizes extracted in this meta-analysis were from studies using serotonergic-system targeting drugs. While this is unsurprising given that commonly prescribed antidepressants target the serotonin system [74], mood disorders are associated with dysfunction in several neurological systems and further investigation of the influence of pharmacologically-induced changes in the activity of these systems may be beneficial [120, 121, 122]. This meta-analysis has highlighted that multiple probe cues may be preferable in future studies. Pharmacological manipulations using affect-altering drugs do not necessarily exert the strongest influence of judgement bias at the most ambiguous cue, as found in this meta-analysis,

and using multiple cues would allow a more comprehensive assessment of the effect of the manipulation. Finally, it would be worthwhile to assess the efficacy of judgement bias as a measure of pharmacological manipulations of affect in mice.

5. Conclusion

To conclude, this meta-analysis has provided evidence that judgement bias has predictive validity as a measure of pharmacologically-induced neurobiological states associated with relatively negative or positive affect, which supports judgement bias as a measure of affect in non-human animals. A key issue identified in this study is the potential interference of drug side effects on judgement bias. In particular, the contrary effect of adrenergic-targeting affect-altering drugs and the greater effect of drugs on judgement bias at lower doses, may be attributed to side effects or to the complex nature of adrenergic drug effects. The effect of depressant and anxiogenic drugs state was greater than the effect of antidepressant and anxiolytic drugs, and therefore larger sample sizes may be required when testing the efficacy of potential pharmacological treatments for mood disorders. However, if consideration is given to these potential shortcomings, the judgement bias task for which there is evidence of construct validity and now of predictive validity appears to be a viable measure of whether the neurobiological state of non-human animals is indicative of a positive or negative affective state.

Acknowledgments

We thank the Biotechnology and Biological Sciences Research Council (BB-SRC: SWBio Doctoral Training Programme grant number BB/M009122/1) for funding this work. SN and ML were supported by an Australian Research Council (ARC) Discovery grant (DP180100818).

References

References

- [1] M. Mendl, O. H. Burman, E. S. Paul, An integrative and functional framework for the study of animal emotion and mood, *Proceedings of the Royal Society of London B: Biological Sciences* 277 (1696) (2010) 2895–2904.
- [2] World Health Organization, *Depression and other common mental disorders: global health estimates* (2017).
- [3] S. S. Valvassori, J. Budni, R. B. Varela, J. Quevedo, Contributions of animal models to the study of mood disorders, *Brazilian Journal of Psychiatry* 35 (2013) S121–S131.
- [4] N. Rupniak, Animal models of depression: challenges from a drug development perspective, *Behavioural pharmacology* 14 (5) (2003) 385–390.
- [5] E. J. Nestler, S. E. Hyman, Animal models of neuropsychiatric disorders, *Nature Neuroscience* 13 (10) (2010) 1161–1169.
- [6] J. De Vry, R. Schreiber, The chronic mild stress depression model: future developments from a drug discovery perspective, *Psychopharmacology* 134 (4) (1997) 349–350.
- [7] R. McArthur, F. Borsini, Animal models of depression in drug discovery: a historical perspective, *Pharmacology Biochemistry and Behavior* 84 (3) (2006) 436–452.
- [8] T. Kato, M. Kubota, T. Kasahara, Animal models of bipolar disorder, *Neuroscience & Biobehavioral Reviews* 31 (6) (2007) 832–842.
- [9] R. Porsolt, M. Le Pichon, M. Jalfre, Depression: a new animal model sensitive to antidepressant treatments, *Nature* 266 (5604) (1977) 730–732.

- [10] J. F. Cryan, A. Markou, I. Lucki, Assessing antidepressant activity in rodents: recent developments and future needs, *Trends in Pharmacological sciences* 23 (5) (2002) 238–245.
- [11] L. Steru, R. Chermat, B. Thierry, P. Simon, The tail suspension test: a new method for screening antidepressants in mice, *Psychopharmacology* 85 (3) (1985) 367–370.
- [12] P. Willner, A. Towell, D. Sampson, S. Sophokleous, R. Muscat, Reduction of sucrose preference by chronic unpredictable mild stress, and its restoration by a tricyclic antidepressant, *Psychopharmacology* 93 (3) (1987) 358–364.
- [13] P. Willner, Validity, reliability and utility of the chronic mild stress model of depression: a 10-year review and evaluation, *Psychopharmacology* 134 (4) (1997) 319–329.
- [14] K. Montgomery, The relation between fear induced by novel stimulation and exploratory drive, *Journal of comparative and physiological psychology* 48 (4) (1955) 254–260.
- [15] S. Pellow, P. Chopin, S. E. File, M. Briley, Validation of open: closed arm entries in an elevated plus-maze as a measure of anxiety in the rat, *Journal of Neuroscience Methods* 14 (3) (1985) 149–167.
- [16] R. Rodgers, B.-J. Cao, A. Dalvi, A. Holmes, Animal models of anxiety: an ethological perspective, *Brazilian Journal of Medical and Biological Research* 30 (1997) 289–304.
- [17] F. Borsini, J. Podhorna, D. Marazziti, Do animal models of anxiety predict anxiolytic-like effects of antidepressants?, *Psychopharmacology* 163 (2) (2002) 121–141.
- [18] A. Carobrez, L. Bertoglio, Ethological and temporal analyses of anxiety-like behavior: the elevated plus-maze model 20 years on, *Neuroscience & Biobehavioral Reviews* 29 (8) (2005) 1193–1205.

- [19] M. L. Molendijk, E. R. de Kloet, Immobility in the forced swim test is adaptive and does not reflect depression, *Psychoneuroendocrinology* 62 (2015) 389–391.
- [20] F. Borsini, A. Meli, Is the forced swimming test a suitable model for revealing antidepressant activity?, *Psychopharmacology* 94 (2) (1988) 147–160.
- [21] A. P. West, Neurobehavioral studies of forced swimming: the role of learning and memory in the forced swim test, *Progress in Neuropsychopharmacology & Biological Psychiatry* 14 (6) (1990) 863–877.
- [22] O. V. Bogdanova, S. Kanekar, K. E. D’Anci, P. F. Renshaw, Factors influencing behavior in the forced swim test, *Physiology & Behavior* 118 (2013) 227–239.
- [23] L. Naudon, T. Jay, Opposite behaviours in the forced swimming test are linked to differences in spatial working memory performances in the rat, *Neuroscience* 130 (2) (2005) 285–293.
- [24] G. S. Dichter, M. J. Smoski, A. B. Kampov-Polevoy, R. Gallop, J. C. Garbutt, Unipolar depression does not moderate responses to the sweet taste test, *Depression and anxiety* 27 (9) (2010) 859–863.
- [25] I. Berlin, L. Givry-Steiner, Y. Lecrubier, A. Puech, Measures of anhedonia and hedonic responses to sucrose in depressive and schizophrenic patients in comparison with healthy subjects, *European Psychiatry* 13 (6) (1998) 303–309.
- [26] N. F. Forbes, C. A. Stewart, K. Matthews, I. C. Reid, Chronic mild stress and sucrose consumption: validity as a model of depression, *Physiology & behavior* 60 (6) (1996) 1481–1484.
- [27] E. J. Harding, E. S. Paul, M. Mendl, Animal behaviour: cognitive bias and affective state, *Nature* 427 (6972) (2004) 312.

- [28] E. J. Bethell, A “how-to” guide for designing judgment bias studies to assess captive animal welfare, *Journal of Applied Animal Welfare Science* 18 (S1) (2015) S18–S42.
- [29] L. Baciadonna, A. G. McElligott, The use of judgement bias to assess welfare in farm livestock, *Animal Welfare* 24 (2015) 81–91.
- [30] T. E. Nygren, A. M. Isen, P. J. Taylor, J. Dulin, The influence of positive affect on the decision rule in risk situations: Focus on outcome (and especially avoidance of loss) rather than probability, *Organizational Behavior and Human Decision Processes* 66 (1) (1996) 59–72.
- [31] E. J. Johnson, A. Tversky, Affect, generalization, and the perception of risk, *Journal of Personality and Social Psychology* 45 (1) (1983) 20–31.
- [32] W. F. Wright, G. H. Bower, Mood effects on subjective probability assessment, *Organizational Behavior and Human Decision Processes* 52 (2) (1992) 276–291.
- [33] E. S. Paul, I. Cuthill, G. Kuroso, V. Norton, J. Woodgate, M. Mendl, Mood and the speed of decisions about anticipated resources and hazards, *Evolution and Human Behavior* 32 (1) (2011) 21–28.
- [34] M. H. Anderson, C. Hardcastle, M. R. Munafò, E. S. Robinson, Evaluation of a novel translational task for assessing emotional biases in different species, *Cognitive, Affective, & Behavioral Neuroscience* 12 (2) (2012) 373–381.
- [35] K. Iigaya, A. Jolivald, W. Jitkrittum, I. D. Gilchrist, P. Dayan, E. Paul, M. Mendl, Cognitive bias in ambiguity judgements: using computational models to dissect the effects of mild mood manipulation in humans, *PloS one* 11 (11) (2016) e0165840.
- [36] N. Kara, Y. Stukalin, H. Einat, Revisiting the validity of the mouse forced swim test: systematic review and meta-analysis of the effects of proto-

typic antidepressants, *Neuroscience & Biobehavioral Reviews* 84 (2017) 1–11.

- [37] R. Yankelevitch-Yahav, M. Franko, A. Huly, R. Doron, The forced swim test as a model of depressive-like behavior, *JoVE (Journal of Visualized Experiments)* (97) (2015) e52587.
- [38] J. F. Cryan, C. Mombereau, A. Vassout, The tail suspension test as a model for assessing antidepressant activity: review of pharmacological and genetic studies in mice, *Neuroscience & Biobehavioral Reviews* 29 (4-5) (2005) 571–625.
- [39] P. Willner, The validity of animal models of depression, *Psychopharmacology* 83 (1) (1984) 1–16.
- [40] M. Mendl, O. H. Burman, R. M. Parker, E. S. Paul, Cognitive bias as an indicator of animal emotion and welfare: emerging evidence and underlying mechanisms, *Applied Animal Behaviour Science* 118 (3-4) (2009) 161–181.
- [41] K. Doya, Modulators of decision making, *Nature neuroscience* 11 (4) (2008) 410.
- [42] N. D. Daw, K. Doya, The computational neurobiology of learning and reward, *Current opinion in neurobiology* 16 (2) (2006) 199–204.
- [43] A. Rangel, C. Camerer, P. R. Montague, A framework for studying the neurobiology of value-based decision making, *Nature Reviews Neuroscience* 9 (7) (2008) 545–556.
- [44] J. A. Burk, S. A. Blumenthal, E. B. Maness, Neuropharmacology of attention, *European journal of pharmacology* 835 (2018) 162–168.
- [45] S. F. Logue, T. J. Gould, The neural and genetic basis of executive function: attention, cognitive flexibility, and response inhibition, *Pharmacology Biochemistry and Behavior* 123 (2014) 45–54.

- [46] M. I. Posner, S. E. Petersen, P. T. Fox, M. E. Raichle, Localization of cognitive operations in the human brain, *Science* 240 (4859) (1988) 1627–1631.
- [47] M. Á. García-Cabezas, B. Rico, M. Á. Sánchez-González, C. Cavada, Distribution of the dopamine innervation in the macaque and human thalamus, *Neuroimage* 34 (3) (2007) 965–984.
- [48] M. J. Owens, C. B. Nemeroff, Role of serotonin in the pathophysiology of depression: focus on the serotonin transporter, *Clinical Chemistry* 40 (2) (1994) 288–295.
- [49] Y.-L. Boureau, P. Dayan, Opponency revisited: competition and cooperation between dopamine and serotonin, *Neuropsychopharmacology* 36 (1) (2011) 74.
- [50] K. C. Berridge, T. E. Robinson, J. W. Aldridge, Dissecting components of reward: ‘liking’, ‘wanting’, and learning, *Current opinion in pharmacology* 9 (1) (2009) 65–73.
- [51] G. Aston-Jones, J. D. Cohen, Adaptive gain and the role of the locus coeruleus–norepinephrine system in optimal performance, *Journal of Comparative Neurology* 493 (1) (2005) 99–110.
- [52] J. Best, M. C. Reed, H. F. Nijhout, Computational studies of the role of serotonin in the basal ganglia, *Frontiers in integrative neuroscience* 7 (2013) 41.
- [53] A. Reiner, L. Medina, C. L. Veenman, Structural and functional evolution of the basal ganglia in vertebrates, *Brain Research Reviews* 28 (3) (1998) 235–285.
- [54] E. Ottaviani, C. Franceschi, The neuroimmunology of stress from invertebrates to man, *Progress in Neurobiology* 48 (4-5) (1996) 421–440.
- [55] J. Venter, U. Di Porzio, D. Robinson, S. Shreeve, J. Lai, A. Kerlavage, S. Frace Jr, K. Lentjes, C. Fraser, Evolution of neurotransmitter receptor systems, *Progress in Neurobiology* 30 (2-3) (1988) 105–169.

- [56] E. M. Jorgensen, Animal evolution: looking for the first nervous system, *Current Biology* 24 (14) (2014) R655–R658.
- [57] S. Ramboz, R. Oosting, D. A. Amara, H. F. Kung, P. Blier, M. Mendelsohn, J. J. Mann, D. Brunner, R. Hen, Serotonin receptor 1a knockout: an animal model of anxiety-related disorder, *Proceedings of the National Academy of Sciences* 95 (24) (1998) 14476–14481.
- [58] E. Bezard, S. Dovero, D. Belin, S. Duconger, V. Jackson-Lewis, S. Przedborski, P. V. Piazza, C. E. Gross, M. Jaber, Enriched environment confers resistance to 1-methyl-4-phenyl-1, 2, 3, 6-tetrahydropyridine and cocaine: involvement of dopamine transporter and trophic factors, *Journal of Neuroscience* 23 (35) (2003) 10999–11007.
- [59] G. Laviola, A. J. Hannan, S. Macrì, M. Solinas, M. Jaber, Effects of enriched environment on animal models of neurodegenerative diseases and psychiatric disorders, *Neurobiology of Disease* 31 (2) (2008) 159–168.
- [60] B. Haenisch, A. Bilkei-Gorzo, M. G. Caron, H. Bönisch, Knockout of the norepinephrine transporter and pharmacologically diverse antidepressants prevent behavioral and brain neurotrophin alterations in two chronic stress models of depression, *Journal of Neurochemistry* 111 (2) (2009) 403–416.
- [61] R. M. Sapolsky, L. M. Romero, A. U. Munck, How do glucocorticoids influence stress responses? integrating permissive, suppressive, stimulatory, and preparative actions, *Endocrine Reviews* 21 (1) (2000) 55–89.
- [62] V. Klimek, C. Stockmeier, J. Overholser, H. Y. Meltzer, S. Kalka, G. Dilley, G. A. Ordway, Reduced levels of norepinephrine transporters in the locus coeruleus in major depression, *Journal of Neuroscience* 17 (21) (1997) 8451–8458.
- [63] J. Massana, Reboxetine versus fluoxetine: an overview of efficacy and tolerability, *The Journal of Clinical Psychiatry* 59 (14) (1998) 8–10.

- [64] G. I. Papakostas, Dopaminergic-based pharmacotherapies for depression, *European Neuropsychopharmacology* 16 (6) (2006) 391–402.
- [65] C. B. Nemeroff, The role of gaba in the pathophysiology and treatment of anxiety disorders, *Psychopharmacology Bulletin* 37 (4) (2003) 133–146.
- [66] A. W. Goddard, G. F. Mason, A. Almai, D. L. Rothman, K. L. Behar, O. A. Petroff, D. S. Charney, J. H. Krystal, Reductions in occipital cortex gaba levels in panic disorder detected with 1h-magnetic resonance spectroscopy, *Archives of General Psychiatry* 58 (6) (2001) 556–561.
- [67] K. A. Dienes, N. A. Hazel, C. L. Hammen, Cortisol secretion in depressed, and at-risk adults, *Psychoneuroendocrinology* 38 (6) (2013) 927–940.
- [68] I. M. Goodyer, A. Bacon, M. Ban, T. Croudace, J. Herbert, Serotonin transporter genotype, morning cortisol and subsequent depression in adolescents, *The British Journal of Psychiatry* 195 (1) (2009) 39–45.
- [69] M. S. D’Souza, Glutamatergic transmission in drug reward: implications for drug addiction, *Frontiers in Neuroscience* 9 (2015) 404.
- [70] R. M. Berman, A. Cappiello, A. Anand, D. A. Oren, G. R. Heninger, D. S. Charney, J. H. Krystal, Antidepressant effects of ketamine in depressed patients, *Biological psychiatry* 47 (4) (2000) 351–354.
- [71] M. Papp, E. Moryl, Antidepressant activity of non-competitive and competitive nmda receptor antagonists in a chronic mild stress model of depression, *European journal of pharmacology* 263 (1-2) (1994) 1–7.
- [72] A. D. Corbett, G. Henderson, A. T. McKnight, S. J. Paterson, 75 years of opioid research: the exciting but vain quest for the holy grail, *British Journal of Pharmacology* 147 (1) (2006) 153–162.
- [73] A. Frasch, Reduction of plasma oxytocin levels in patients suffering from major depression, *Adv Exp Med Biol* 395 (1995) 257–258.

- [74] B. Mars, J. Heron, D. Gunnell, R. M. Martin, K. H. Thomas, D. Kessler, Prevalence and patterns of antidepressant switching amongst primary care patients in the uk, *Journal of Psychopharmacology* 31 (5) (2017) 553–560.
- [75] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, Preferred reporting items for systematic reviews and meta-analyses: the prisma statement, *Annals of Internal Medicine* 151 (4) (2009) 264–269.
- [76] D. S. Wishart, Y. D. Feunang, A. C. Guo, E. J. Lo, A. Marcu, J. R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, et al., Drugbank 5.0: a major update to the drugbank database for 2018, *Nucleic Acids Research* 46 (D1) (2017) D1074–D1082.
- [77] M. E. Thase, Effects of venlafaxine on blood pressure: a meta-analysis of original data from 3744 depressed patients., *The Journal of Clinical Psychiatry* 59(10) (1998) 502–508.
- [78] C. A. Shaheed, C. G. Maher, K. A. Williams, R. Day, A. J. McLachlan, Efficacy, tolerability, and dose-dependent effects of opioid analgesics for low back pain: a systematic review and meta-analysis, *JAMA internal medicine* 176 (7) (2016) 958–968.
- [79] S. Nakagawa, P. C. Johnson, H. Schielzeth, The coefficient of determination r^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded, *Journal of the Royal Society Interface* 14 (134) (2017) 20170213.
- [80] S. Nakagawa, H. Schielzeth, Repeatability for gaussian and non-gaussian data: a practical guide for biologists, *Biological Reviews* 85 (4) (2010) 935–956.
- [81] L. V. Hedges, Distribution theory for glass's estimator of effect size and related estimators, *journal of Educational Statistics* 6 (2) (1981) 107–128.

- [82] J. P. Higgins, S. Green, Cochrane handbook for systematic reviews of interventions 5.1. 0, The Cochrane Collaboration (2011) 33–49.
- [83] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2017).
URL <http://www.R-project.org/>
- [84] W. Viechtbauer, Conducting meta-analyses in R with the metafor package, *Journal of Statistical Software* 36 (3) (2010) 1–48.
URL <http://www.jstatsoft.org/v36/i03/>
- [85] S. Nakagawa, D. W. Noble, A. M. Senior, M. Lagisz, Meta-evaluation of meta-analysis: ten appraisal questions for biologists, *BMC biology* 15 (1) (2017) 18.
- [86] W. Van den Noortgate, J. A. López-López, F. Marín-Martínez, J. Sánchez-Meca, Three-level meta-analysis of dependent effect sizes, *Behavior research methods* 45 (2) (2013) 576–594.
- [87] G. Knapp, J. Hartung, Improved tests for a random effects meta-regression with a single covariate, *Statistics in medicine* 22 (17) (2003) 2693–2710.
- [88] S. Nakagawa, E. S. Santos, Methodological issues and advances in biological meta-analysis, *Evolutionary Ecology* 26 (5) (2012) 1253–1274.
- [89] J. P. Higgins, S. G. Thompson, Quantifying heterogeneity in a meta-analysis, *Statistics in medicine* 21 (11) (2002) 1539–1558.
- [90] W. Viechtbauer, Conducting meta-analyses in r with the metafor package, *Journal of statistical software* 36 (3) (2010) 1–48.
- [91] J. D. Hadfield, Mcmc methods for multi-response generalized linear mixed models: The MCMCglmm R package, *Journal of Statistical Software* 33 (2) (2010) 1–22.
URL <http://www.jstatsoft.org/v33/i02/>

- [92] M. Egger, G. D. Smith, M. Schneider, C. Minder, Bias in meta-analysis detected by a simple, graphical test, *Bmj* 315 (7109) (1997) 629–634.
- [93] J. Cohen, *Statistical power analysis for the behavioral sciences* 2nd edn (1988).
- [94] S. Nakagawa, H. Schielzeth, A general and simple method for obtaining r^2 from generalized linear mixed-effects models, *Methods in Ecology and Evolution* 4 (2) (2013) 133–142.
- [95] M. Tanaka, M. Yoshida, H. Emoto, H. Ishii, Noradrenaline systems in the hypothalamus, amygdala and locus coeruleus are involved in the provocation of anxiety: basic studies, *European Journal of Pharmacology* 405 (1-3) (2000) 397–406.
- [96] T. Inoue, S. Nakagawa, T. Izumi, Y. Kitaichi, T. Koyama, Effect of combined treatment with noradrenaline and serotonin reuptake inhibitors on conditioned freezing, *European Journal of Pharmacology* 540 (1-3) (2006) 91–95.
- [97] C. Harmer, J. Heinzen, U. O’Sullivan, R. Ayres, P. Cowen, Dissociable effects of acute antidepressant drug administration on subjective and emotional processing measures in healthy volunteers, *Psychopharmacology* 199 (4) (2008) 495–502.
- [98] D. A. Morilak, G. Barrera, D. J. Echevarria, A. S. Garcia, A. Hernandez, S. Ma, C. O. Petre, Role of brain norepinephrine in the behavioral response to stress, *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 29 (8) (2005) 1214–1224.
- [99] K. A. Roth, I. M. Mefford, J. D. Barchas, Epinephrine, norepinephrine, dopamine and serotonin: differential effects of acute and chronic stress on regional brain amines, *Brain Research* 239 (2) (1982) 417–424.
- [100] A. M. Senior, C. E. Grueber, T. Kamiya, M. Lagisz, K. O’Dwyer, E. S. Santos, S. Nakagawa, Heterogeneity in ecological and evolutionary meta-

analyses: its magnitude and implications, *Ecology* 97 (12) (2016) 3293–3299.

- [101] E. J. Bethell, N. F. Koyama, Happy hamsters? enrichment induces positive judgement bias for mildly (but not truly) ambiguous cues to reward and punishment in *Mesocricetus auratus*, *Royal Society open science* 2 (7) (2015) 140399.
- [102] Center for Drug Evaluation and Research, FDA Adverse Event Reporting System (FAERS) Public Dashboard.
URL <https://www.fda.gov/drugs/guidancecomplianceregulatoryinformation/surveillance/adversedrugs/effects/ucm070093.htm>
- [103] J. Kukulja, T. E. Schläpfer, C. Keyzers, D. Klingmüller, W. Maier, G. R. Fink, R. Hurlemann, Modeling a negative response bias in the human amygdala by noradrenergic–glucocorticoid interactions, *Journal of neuroscience* 28 (48) (2008) 12868–12876.
- [104] H. Buerkle, T. Yaksh, Pharmacological evidence for different alpha 2-adrenergic receptor sites mediating analgesia and sedation in the rat, *British Journal of Anaesthesia* 81 (2) (1998) 208–215.
- [105] A. C. Parrott, A. Gibbs, A. B. Scholey, R. King, K. Owens, P. Swann, E. Ogden, C. Stough, Mdma and methamphetamine: some paradoxical negative and positive mood changes in an acute dose laboratory study, *Psychopharmacology* 215 (3) (2011) 527–536.
- [106] M.-L. Foisy, C. Kornreich, A. Fobe, L. D’hondt, I. Pelc, C. Hanak, P. Verbanck, P. Philippot, Impaired emotional facial expression recognition in alcohol dependence: do these deficits persist with midterm abstinence?, *Alcoholism: Clinical and Experimental Research* 31 (3) (2007) 404–410.
- [107] R. M. Nesse, The smoke detector principle, *Annals of the New York Academy of Sciences* 935 (1) (2001) 75–85.

- [108] D. Nettle, M. Bateson, The evolutionary origins of mood and its disorders, *Current Biology* 22 (17) (2012) R712–R721.
- [109] E. J. Calabrese, L. A. Baldwin, U-shaped dose-responses in biology, toxicology, and public health, *Annual review of public health* 22 (1) (2001) 15–33.
- [110] E. J. Calabrese, L. A. Baldwin, Hormesis: U-shaped dose responses and their centrality in toxicology, *Trends in pharmacological sciences* 22 (6) (2001) 285–291.
- [111] E. J. Calabrese, L. A. Baldwin, The frequency of u-shaped dose responses in the toxicological literature, *Toxicological Sciences* 62 (2) (2001) 330–338.
- [112] J. Lazarou, B. H. Pomeranz, P. N. Corey, Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies, *JAMA* 279 (15) (1998) 1200–1205.
- [113] T. A. Furukawa, H. McGuire, C. Barbui, Meta-analysis of effects and side effects of low dosage tricyclic antidepressants in depression: systematic review, *BMJ* 325 (2002) 991–995.
- [114] M. Mendl, E. S. Paul, P. Dayan, Integrating animal affect and decision-making: a reinforcement-learning perspective (in prep.).
- [115] P. Trimmer, E. Paul, M. Mendl, J. McNamara, A. Houston, On the evolution and optimality of mood states, *Behavioral Sciences* 3 (3) (2013) 501–521.
- [116] Home Office, UK, Annual statistics of scientific procedures on living animals, Great Britain 2016 (2017).
- [117] National Research Council, Use of laboratory animals in biomedical and behavioral research, National Academies Press, 1988.

- [118] S. Hintze, L. Melotti, S. Colosio, J. D. Bailoo, M. Boada-Saña, H. Würbel, E. Murphy, A cross-species judgement bias task: integrating active trial initiation into a spatial go/no-go task, *Scientific reports* 8 (1) (2018) 5104.
- [119] V. Krakenberg, I. Woigk, L. G. Rodriguez, N. Kästner, S. Kaiser, N. Sachser, S. H. Richter, Technology or ecology? new tools to assess cognitive judgement bias in mice, *Behavioural brain research* 362 (2019) 279–287.
- [120] C. B. Nemeroff, M. J. Owens, Treatment of mood disorders, *Nature Neuroscience* 5 (2002) 1068–1070.
- [121] E. J. Nestler, M. Barrot, R. J. DiLeone, A. J. Eisch, S. J. Gold, L. M. Monteggia, Neurobiology of depression, *Neuron* 34 (1) (2002) 13–25.
- [122] M. Joëls, T. Z. Baram, The neuro-symphony of stress, *Nature Reviews Neuroscience* 10 (6) (2009) 459–466.

Figure Captions

Figure 1: PRISMA Flow Diagram illustrating the number (n) of articles included at each stage of the literature review.

Figure 2: Example of hypothesised data from the judgement bias task with two treatments; one designed to induce a relatively positive affective state (relatively favourable treatment) and another designed to induce a relatively negative affective state (relatively unfavourable treatment). While the mean proportion of positive responses is almost identical at the positive and negative reference cue, a treatment difference is observed at the probe cues.

Figure 3: Forest plot with a meta-analytic mean (intercept-only model) and significant moderators from univariate meta-regression models. Each point represents the mean effect size for each moderator and error bars represent the 95% confidence interval.

Figure 4: Forest plot with a meta-analytic mean (intercept-only model) and significant moderators from univariate meta-regression models following the exclusion of adrenergic system targeting drugs. Each point represents the mean effect size for each moderator and error bars represent the 95% confidence interval.

Figure 5: Funnel plots of a) the meta-analytic residual values (residuals + sampling errors) for the full meta-regression model prior to exclusion of effect sizes from studies using adrenergic system targeting drugs; b) the raw effect sizes and the inverse standard errors prior to exclusion of effect sizes from studies using adrenergic system targeting drugs; c) the meta-analytic residual values for the full meta-regression model following exclusion of effect sizes from studies using adrenergic system targeting drugs; d) the raw effect sizes and the inverse standard errors following exclusion of effect sizes from studies using adrenergic system targeting drugs.

Figures

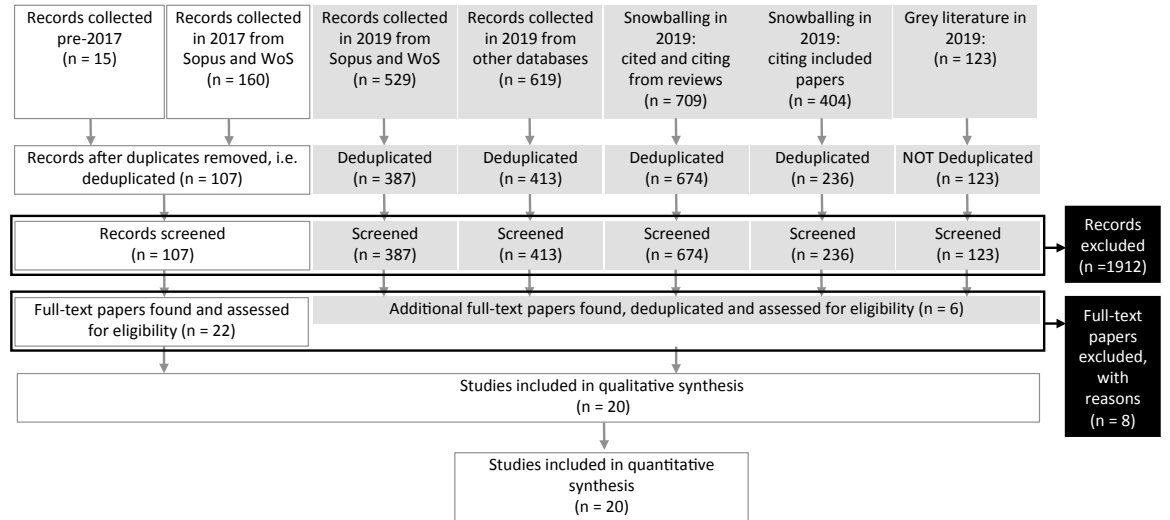


Figure 1

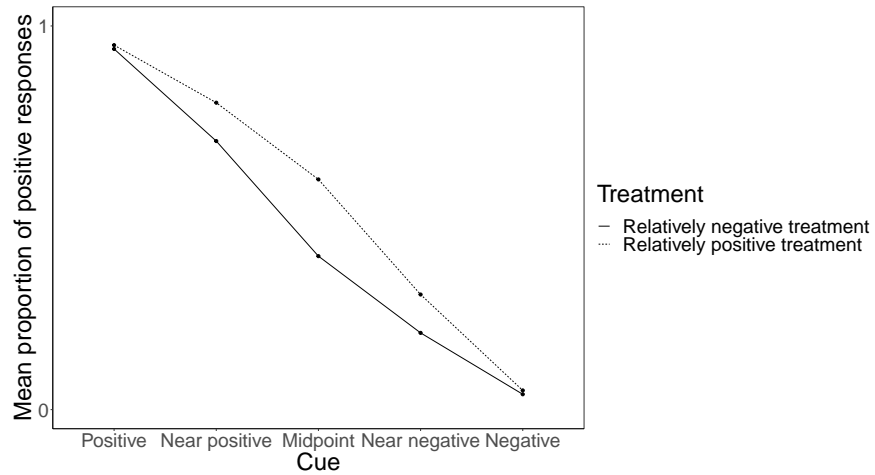


Figure 2

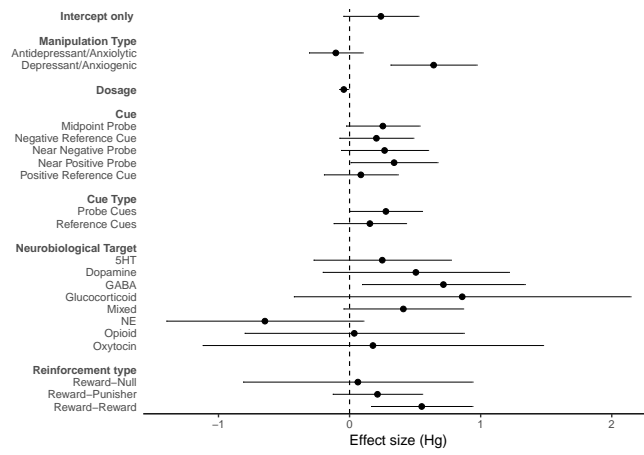


Figure 3

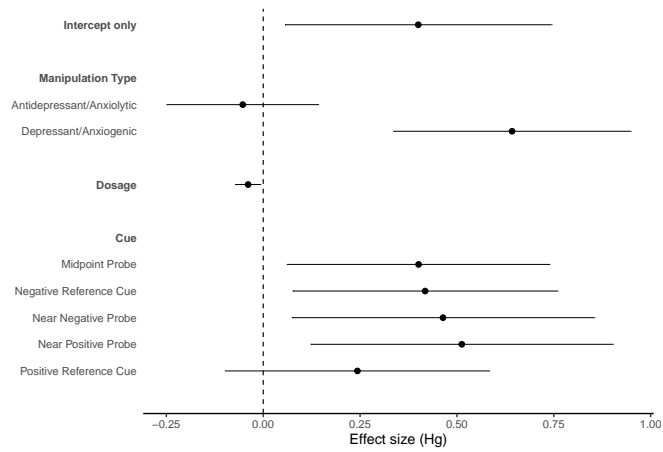


Figure 4

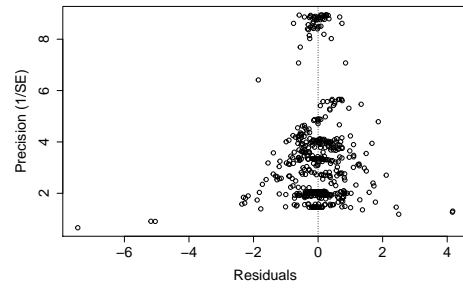


Figure 5a

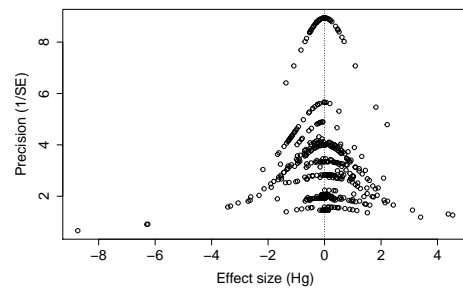


Figure 5b

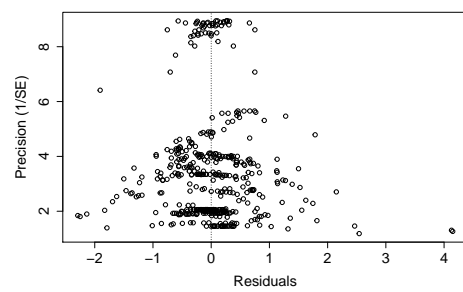


Figure 5c

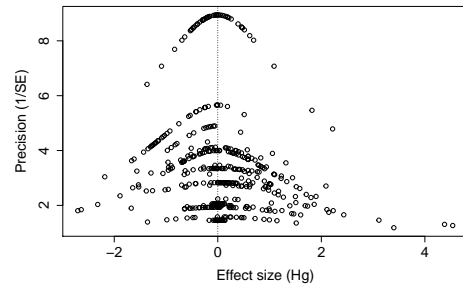


Figure 5d

Tables

Table 1: Articles included at full-text screening and reason for exclusion, where relevant

Article number	Status	Authors	Article title	Journal	Year	Reason for exclusion
1	Included	Anderson, M.H., Munafò, M.R., Robinson, E.S.J.	Investigating the psychopharmacology of cognitive affective bias in rats using an affective tone discrimination task	Psychopharmacology	2013	NA
2	Included	Destrez, A., Deiss, V., Belzung, C., Lee, C., Bussey, A.	Does reduction of fearfulness tend to reduce pessimistic-like judgment in lambs?	Applied Animal Behaviour Science	2012	NA
3	Included	Doyle, R.E., Hinch, G.N., Fisher, A.D., Bussey, A., Henshall, J.M., Lee, C.	Administration of serotonergic inhibitor p-Chlorophenylalanine induces pessimistic-like judgment bias in sheep	Psychoneuroendocrinology	2011	NA
4	Included	Eikel, T., Ghelardath, D., Van Bollen Und Hülbach, O., Sanchez-Segura, C., Hurlmann, R., Spanagel, R., Gass, P., Wilmshäfer, B.	Ambiguous cue interpretation is biased under stress and depression-like states in rats	Neuropsychopharmacology	2010	NA
5	Included	Golebiowska, G., Rygula, R.	Effects of acute dopaminergic and serotonergic manipulations in the ACI paradigm depend on the basal valence of cognitive judgment bias in rats	Behavioural Brain Research	2017	NA
6	Included	Hales, C.A., Robinson, E.S.J., Houghton, C.J.	Diffusion modelling reveals the decision making processes underlying negative judgment bias in rats	PLoS One	2016	NA
7	Included	Hales, C.A., Houghton, C.J., Robinson, E.S.J.	Behavioural and computational methods reveal differential effects for how delayed and rapid onset antidepressants affect decision making in rats	European Neuropsychopharmacology	2017	NA
8	Included	Hymel, K.A., Sufka, K.J.	Pharmacological reversal of cognitive bias in the chick anxiety-depression model	Neuropharmacology	2010	NA
9	Included	Ilyasene, O.S., Beard, A.P., Guy, H., Bateson, M.	Elevated levels of the stress hormone, corticosterone, cause "pessimistic" judgment bias in broiler chickens	Scientific Reports	2017	NA
10	Included	Kis, A., Herr di, A., Kanizsa, T.O., Gasi, M., Top, I., J.	Oxytocin induces positive expectations about ambivalent stimuli (cognitive bias) in dogs	Hormones and Behavior	2015	NA
11	Included	McGuire, M.C., Williams, K.L., Wellings, L.L.M., York, J.	Cognitive bias in rats is not influenced by oxytocin	Frontiers in Psychology	2015	NA
12	Included	Rygula, R., Golebiowska, J., Krugiel, J., Holst, M., Popik, P.	Acute administration of lithium, but not valproate, modulates cognitive judgment bias in rats	Psychopharmacology	2015	NA
13	Included	Rygula, R., Popik, J., Popik, P.	The effects of acute pharmacological stimulation of the 5-HT _{1A} and DA systems on the cognitive judgment bias of rats in the ambiguous-cue interpretation paradigm	European Neuropsychopharmacology	2014	NA
14	Included	Rygula, R., Szczeciński, E., Krugiel, J., Golebiowska, J., Kubiś, J., Popik, P.	Cognitive judgment bias in the psychostimulant-induced model of mania in rats	Psychopharmacology	2015	NA
15	Included	Rygula, R., Szczeciński, E., Popik, J., Nikiforuk, A., Popik, P.	The effects of cocaine and meprobamate on the cognitive judgment bias of rats in the ambiguous-cue interpretation paradigm	Behavioural Brain Research	2014	NA
16	Included	Sahin, C., Dowdall, N., Neill, J.C.	Towards the development of improved tests for negative symptoms of schizophrenia in a validated animal model	Behavioural Brain Research	2016	NA
17	Included	Stracke, J., Otten, W., Tüschner, A., Purpke, B., Dupjan, S.	Serotonin depletion induces pessimistic-like behavior in a cognitive bias paradigm in pigs	Physiology and Behavior	2017	NA
18	Included	Stracke, J., Otten, W., Tüschner, A., Withalm, M., Metzger, C.C., Purpke, B., Dupjan, S.	Dietary tryptophan supplementation and affective state in pigs	Journal of Veterinary Behaviour	2017	NA
19	Included	Verbeek, E., Ferguson, D., Lee, C.	Are hungry sheep more pessimistic? The effects of food restriction on cognitive bias and the involvement of ghrelin in its regulation	Physiology and Behavior	2014	NA
20	Included	Verbeek, E., Ferguson, D., Quiquet de Marjoux, P., Lee, C.	Generating positive affective states in sheep: The influence of food rewards and opioid administration	Applied Animal Behaviour Science	2014	NA
21	Not included	Anderson, M.H., Munafò, M.R., Robinson, E.S.J.	The effects of acute psychopharmacological treatments on cognitive affective bias in rats	European Neuropsychopharmacology	2012	Conference abstract that duplicates Anderson et al 2013
22	Not included	Hales, C., Bartlett, J., Aban, R., Hengeler, B., Robinson, E.	Targeted infusions with rapid acting antidepressants reveal a role for the prefrontal cortex in mediating affective biases and decision making	Brain and Neuroscience Advances	2019	Data not available
23	Not included	Karagiannis, C.L., Burman, O.H.P., Mills, D.S.	Dogs with separation-related problems show a "less pessimistic" cognitive bias during treatment with fluoxetine (Reconcile [®]) and a behaviour modification plan	BMC Veterinary Research	2015	Data not available
24	Not included	Krugiel, J., Golebiowska, J., Popik, P., Rygula, R.	Dopamine induces an optimism bias in rats: pharmacological proof for the translational validity of the ambiguous-cue interpretation test	Behavioural Brain Research	2016	Retracted by author
25	Not included	Krugiel, J., Males, N., Popik, P., Starowicz, K., Rygula, R.	Anandamide mediates cognitive judgment bias in rats	Neuropharmacology	2016	Retracted by author
26	Not included	Neill, J., Gabriel, W., Wolter, W., Teitler, V.	NMDA receptor antagonists in rodents, relevance to negative symptoms of schizophrenia: A translational link to humans	European Archives of Psychiatry and Clinical Neuroscience	2015	Not experimental research article
27	Not included	Phillips B.U., Dewan, S., Nilsson, S.R.O., Robbins, T.W., Heath, C.J., Saksida L.M., Bussey, T.J., Also, J.	Selective effects of 5-HT _{2C} receptor modulation on performance of a novel valence probe visual discrimination task and probabilistic reversal learning in mice	Psychopharmacology	2018	Didn't use a variant of Harding et al's task
28	Not included	Sahin, C., Podda, G., Grayson, B., Marsh, S., Aricioglu, F., Neill, J.C.	The deficit in anticipatory motivation as a negative symptom of schizophrenia: Phenylcyclidine treated rats exhibit pessimism in an optimistic bias task	European Neuropsychopharmacology	2015	Conference abstract that duplicates Sahin et al 2016

Table 2: Table 2: Information extracted from each article included in the analysis

Article number	Drug	Pharmacological target	Doses	Dosing frequency	Time between administration and testing	Number of administration days prior to testing	Number of administration days between final treatment and testing	Manipulation	Species	Sex	Response type	Reinforcement type	Outcome variable	Number of probe cues	Proportion of probe cues
1	diazepam	GABergic system	0, 0.3, 1	acute	30	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 1	1	0.3
	fluoxetine	serotonergic system	0, 1	chronic (wash-out)	NA	1,4,8,11,15,18	NA	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 1	1	0.3
	fluoxetine	serotonergic system	0, 0.3, 1, 3	acute	60	NA	5,7	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 1	1	0.3
	fluoxetine	serotonergic system	0, 0.1, 0.3, 1	acute	30	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 3	3	0.3
2	reboxetine	adrennergic system	0, 0.1, 0.3, 1	acute	30	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 3	3	0.3
	diazepam	GABergic system	0, 0.1	acute	10,180	NA	NA	antidepressant / anisolytic	sheep	female	go/no-go	reward/punisher	latency	3	0.6
3	P-Chlorophenylalanine	serotonergic system	0, 40	chronic	NA	3,5	NA	depressant / anisogenic	sheep	female	go/no-go	reward/punisher	proportion 3	3	0.6
	P-Chlorophenylalanine	serotonergic system	0, 40	chronic (wash-out)	NA	NA	5	depressant / anisogenic	sheep	female	go/no-go	reward/punisher	proportion 3	3	0.6
4	corticosterone+HBC complex + reboxetine	multiple	0, 0.5 (cont) + 15 (hbc)	acute	30 (cont) + 60 (hbc)	NA	NA	depressant / anisogenic	rat	male	go/go	reward/punisher	proportion 3	3	0.25
5	escitalopram	serotonergic system	0, 0.5, 1, 2	acute	30	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 1	1	0.2
	haloperidol	depannergic system	0, 0.01, 0.02, 0.05	acute	30	NA	NA	depressant / anisogenic	rat	male	go/go	reward/punisher	proportion 1	1	0.2
	l-dopa	depannergic system	0, 2, 4, 8	acute	30	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 1	1	0.2
6	FZ7142	GABergic system	0, 3.0, 5.0	acute	30	NA	NA	depressant / anisogenic	rat	male	go/go	reward/reward	proportion 1	1	0.3
7	fluoxetine	serotonergic system	0, 0.3, 1	acute	30	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/reward	proportion 1	1	0.3
	fluoxetine	serotonergic system	0, 1	chronic	NA	1,4,8,11,15,18	NA	antidepressant / anisolytic	rat	male	go/go	reward/reward	proportion 1	1	0.3
	reboxetine	serotonergic system	0, 0.3, 1	acute (wash-out)	30	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/reward	proportion 1	1	0.3
	venlafaxine	adrennergic system	0, 1, 3	acute	30	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/reward	proportion 1	1	0.3
	ketamine	multiple	0, 0.3, 1, 3	acute	60	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/reward	proportion 1	1	0.3
	phenylethylamine	multiple	0, 0.3, 1, 3	acute	40	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/reward	proportion 1	1	0.3
	amphetamine	multiple	0, 0.1, 0.3	acute	15	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/reward	proportion 1	1	0.3
	cocaine	multiple	0, 0.3, 1, 3	acute	10	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/reward	proportion 1	1	0.3
8	imipramine	multiple	0, 15	acute	15	NA	NA	antidepressant / anisolytic	chicken	male	go/no-go	reward/punisher	latency	2	0.5
	clonidine	adrennergic system	0, 0.1	acute	15	NA	NA	antidepressant / anisolytic	chicken	male	go/no-go	reward/punisher	latency	2	0.5
9	corticosterone	glucocorticoid system	0, 4	chronic	NA	3,4,5	NA	depressant / anisogenic	chicken	female	go/no-go	reward/punisher	latency	3	0.3
10	oxytocin	oxytocin system	NA	acute	40	NA	NA	antidepressant / anisolytic	dog	mixed	go/no-go	reward/null	latency	1	0.3
11	oxytocin	oxytocin system	0, 0.001	acute	5	NA	NA	antidepressant / anisolytic	rat	male	go/no-go	reward/punisher	latency	1	0.25
12	lithium chloride	multiple	0, 10, 50, 100	acute	30	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 1	1	0.2
	valproic acid	GABergic system	0, 100, 200, 400	acute	30	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 1	1	0.2
13	clitalopram	serotonergic system	0, 1, 5, 10	acute	30	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 1	1	0.2
	d-amphetamine	multiple	0, 0.1, 0.5, 1	acute	30	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 1	1	0.2
	desipramine	multiple	0, 1, 2, 5	acute	30	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 1	1	0.2
14	cocaine	multiple	0, 10	chronic	NA	14	NA	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 1	1	0.2
	d-amphetamine	multiple	0, 2	chronic	NA	14	NA	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 1	1	0.2
15	cocaine	multiple	0, 1, 2, 5	acute	30	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 1	1	0.2
	mazindol	multiple	0, 0.5, 1, 2	acute	30	NA	NA	antidepressant / anisolytic	rat	male	go/go	reward/punisher	proportion 1	1	0.2
16	phenylethylamine	multiple	0, 2	chronic (wash-out)	NA	NA	8, 9, 10, 11, 12	depressant / anisogenic	rat	female	go/go	reward/reward	proportion 1	1	0.5
17	P-Chlorophenylalanine	serotonergic system	0, 50	chronic (wash-out)	NA	NA	1,2,3,5,10	depressant / anisogenic	pig	female	go/no-go	reward/punisher	latency	3	0.14
18	tryptophan	serotonergic system	NA	chronic	NA	6,7,8,13,14,15	NA	antidepressant / anisolytic	sheep	female	go/no-go	reward/punisher	proportion 3	3	0.14
19	ghrelin	multiple	0, 0.007	acute	10	NA	NA	depressant / anisogenic	sheep	female	go/no-go	reward/punisher	proportion 3	3	0.6
20	morphine	opioid system	0, 1	acute	10	NA	NA	antidepressant / anisolytic	sheep	female	go/no-go	reward/punisher	latency	3	0.6
	naltrexone	opioid system	0, 2	acute	10	NA	NA	depressant / anisogenic	sheep	female	go/no-go	reward/punisher	latency	3	0.6

Table 3: Pairwise comparison of each level of significant moderators from the meta-regression

Variable	Model	Mean difference	CI lower bound	CI upper bound	p-value
<i>Cue</i>	Midpoint - Positive	0.028	0.307	2.36	0.019
	Negative - Positive	-0.022	0.26	1.652	0.099
	Near Negative - Positive	-0.059	0.421	1.479	0.140
	Near Positive - Positive	0.015	0.492	2.087	0.037
	Midpoint - Near Positive	-0.324	0.153	-0.707	0.480
	Negative - Near Positive	-0.375	0.105	-1.105	0.270
	Near Negative - Near Positive	-0.357	0.212	-0.501	0.616
	Negative - Midpoint	-0.190	0.091	-0.686	0.493
	Near Negative - Midpoint	-0.227	0.253	0.108	0.914
	Negative - Near Negative	-0.304	0.179	-0.507	0.612
<i>Cue type</i>	Reference - Probe	-0.232	-0.012	-2.174	0.030
<i>Manipulation Type</i>	Negative - Positive	0.38	1.112	4.007	<0.001
<i>Neurobiological Target</i>	Serotonergic - Adrenergic	0.118	1.672	2.262	0.024
	Dopaminergic - Adrenergic	0.254	2.049	2.521	0.012
	GABAergic - Adrenergic	0.54	2.184	3.254	0.001
	Glucocorticoid - Adrenergic	0.017	2.993	1.987	0.047
	Multiple - Adrenergic	0.346	1.766	2.922	0.004
	Opioid - Adrenergic	-0.365	1.729	1.279	0.201
	Oxytocin - Adrenergic	-0.676	2.324	1.079	0.281
	Serotonergic - Multiple	-0.622	0.3	-0.685	0.493
	Dopaminergic - Multiple	-0.533	0.724	0.298	0.766
	GABAergic - Multiple	-0.244	0.855	1.092	0.276
	Glucocorticoid - Multiple	-0.914	1.812	0.647	0.518
	Opioid - Multiple	-1.208	0.459	-0.882	0.378
	Oxytocin - Multiple	-1.608	1.144	-0.331	0.741
	Dopaminergic - Serotonergic	-0.435	0.948	0.729	0.467
	GABAergic - Serotonergic	-0.156	1.088	1.473	0.141
	Glucocorticoid - Serotonergic	-0.777	1.996	0.864	0.388
	Opioid - Serotonergic	-1.035	0.608	-0.511	0.610
	Oxytocin - Serotonergic	-1.471	1.328	-0.1	0.920
	GABAergic - Dopaminergic	-0.542	0.962	0.548	0.584
	Glucocorticoid - Dopaminergic	-1.114	1.821	0.473	0.636
	Opioid - Dopaminergic	-1.469	0.529	-0.924	0.356
	Oxytocin - Dopaminergic	-1.807	1.152	-0.435	0.664
	Glucocorticoid - GABAergic	-1.283	1.570	0.198	0.843
	Opioid - GABAergic	-1.626	0.266	-1.411	0.159
	Oxytocin - GABAergic	-1.976	0.901	-0.734	0.463
	Opioid - Glucocorticoid	-2.355	0.708	-1.056	0.292
	Oxytocin - Glucocorticoid	-2.507	1.145	-0.733	0.464
	Oxytocin - Opioid	-1.401	1.686	0.181	0.856
<i>Reinforcement Type</i>	Reward/Punisher - Reward/Null	-0.702	1.002	0.345	0.730
	Reward/Reward - Reward/Null	-0.395	1.369	1.085	0.279
	Reward/Reward - Reward/Punisher	0.089	0.585	2.673	0.008